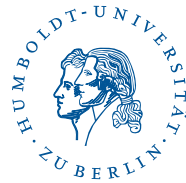


HUMBOLDT-UNIVERSITÄT ZU BERLIN  
WIRTSCHAFTSWISSENSCHAFTLICHE FAKULTÄT  
LADISLAUS VON BORTKIEWICZ CHAIR OF STATISTICS



ANWENDUNG DER METHODEN UND MODELLE DER  
ZEITREIHENANALYSE AUF SPOT- UND  
KRAFTSTOFFPREISE

APPLYING TIME SERIES METHODS AND MODELS TO SPOT AND GASOLINE PRICES



BACHELORARBEIT

VON

IRYNA KARPENKA

ZUR ERLANGUNG DES AKADEMISCHEN GRADES

BACHELOR IN SCIENCE (B.Sc.)

IN BETRIEBSWIRTSCHAFTSLEHRE

MATRIKEL-Nr. 513416

PRÜFER: PROF. DR. WOLFGANG HÄRDLE

PROF. DR. OSTAP OKHRIN

BETREUER: DR. SIGBERT KLINKE

8.01.2012, BERLIN

# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>2</b>
<b>2</b>	<b>Analytische Grundlagen und Methoden</b>	<b>4</b>
2.1	Zeitreihen und Zeitreihenmodelle . . . . .	4
2.2	Stationarität . . . . .	5
2.2.1	Deterministische und stochastische Trends . . . . .	5
2.2.2	Einheitswurzel-Tests . . . . .	7
2.2.3	Residuen-Analyse . . . . .	9
2.3	Saisonalität . . . . .	10
2.3.1	Saisonbereinigung von Zeitreihen . . . . .	11
2.3.2	Modelle für saisonale Daten . . . . .	13
2.4	Transformation von Zeitreihen durch Filter . . . . .	13
2.4.1	Methode der gleitenden Durchschnitte . . . . .	13
2.4.2	Methode des Differenzenfilters . . . . .	15
2.5	Modellschätzung . . . . .	15
2.5.1	ARMA-Modelle . . . . .	15
2.5.2	AR(p)-Prozess . . . . .	16
2.5.3	MA-Prozess . . . . .	17
<b>3</b>	<b>Analyse der Zeitreihen auf Stationarität</b>	<b>18</b>
3.1	Vorbereitung der Daten . . . . .	18
3.2	Vorläufige Analyse . . . . .	18
3.3	Einheitswurzel-Tests auf Stationarität . . . . .	19
3.3.1	Erweiterter Dickey-Fuller-Test . . . . .	19
3.3.2	Kwiatkowski-Phillips-Schmidt-Shin-Test . . . . .	22
<b>4</b>	<b>Periodizität und Saisonalität</b>	<b>25</b>
4.1	Analyse der Zeitreihen auf Periodizität . . . . .	25
4.2	Saisonale Zerlegung der Zeitreihen . . . . .	31
4.2.1	Multiplikatives und additives Modell . . . . .	32
4.2.2	Anwendung des additiven Modells . . . . .	36
<b>5</b>	<b>Modellschätzung</b>	<b>39</b>
5.1	Modellbildung der Spotpreise . . . . .	39
5.2	Modellbildung der Dieselpreise . . . . .	42
5.3	Korrelation zwischen den Kraftstoffpreisen und den Spotpreisen . . . . .	44
<b>6</b>	<b>Zusammenfassung</b>	<b>48</b>
<b>7</b>	<b>Literaturverzeichnis</b>	<b>50</b>
<b>8</b>	<b>Anhang</b>	<b>52</b>

# 1 Einführung

Es wird in der Presse oft diskutiert, ob und wie die Spotpreise die Kraftstoffpreise an den Tankstellen beeinflussen. Man merkt oft, wie zum Wochenende die Kraftstoffpreise an den Tankstellen steigen und danach wieder fallen. Zu betonen ist auch die Tatsache, dass die Preise am Wochenende im Durchschnitt höher als werktags sind. In diesem Zusammenhang wäre es für den Endverbraucher sinnvoll, im Voraus zu planen, an welchen Tagen man am besten tanken wird. Die Schuld an solchem Verhalten der Preise wird oft den Tankstellen zugeschrieben, weil sie den Endpreis des an den Konsumenten zu verkaufenden Kraftstoffes bestimmen.

Die kurzfristige Änderung des Kraftstoffpreises an der Zapfsäule hängt von den Tankstellen selbst ab, das langfristige Verhalten wird aber durch den Welterdölpreis und die fünf großen Mineralölunternehmen BP (Aral), ConocoPhillips (Jet), ExxonMobil (Esso), Shell und Total bestimmt. Die Macht dieser fünf großen Konzerne in Deutschland hat in den letzten Jahren sehr zugenommen. Das Bundeskartellamt hat eine umfangreiche Recherche durchführen lassen. In dem am 26. Mai 2011 veröffentlichten „Abschlussbericht zur Sektoruntersuchung der Kraftstoffe“ wurde eine eingehende Analyse der Wettbewerbsverhältnisse auf den Tankstellenmärkten in Deutschland durchgeführt. Der Präsident des Bundeskartellamtes Andreas Mundt hat dazu folgende Stellung bezogen: „Die fünf großen Tankstellenbetreiber in Deutschland machen sich gegenseitig keinen wesentlichen Wettbewerb, sie bilden ein marktbeherrschendes Oligopol. Unsere Studie weist im Einzelnen nach, wie die Mechanismen der Preissetzung funktionieren. Es bedarf bei solchen Marktstrukturen nicht zwingend einer Absprache. Die Unternehmen verstehen sich ohne Worte. Das führt zu überhöhten Preisen“ (Bundeskartellamt, 2011).

Das Ziel dieser Arbeit besteht in dem Aufzeigen vom Verhalten der Kraftstoffpreise und deren Zusammenhangs mit den Spotpreisen anhand der Zeitreihenanalysemethoden und mittels der spezialisierten Software (Excel und Eviews) zu analysieren. Da die zur Verfügung stehenden empirischen Daten während eines Zeitraums von etwas mehr als einem Jahr erhoben wurden, wird mehr auf das wöchentliche Verhalten der Kraftstoffpreise eingegangen. Es wird nach dem Modell gesucht, das optimalerweise die empirischen Daten beschreibt und eine gut geeignete Prognose liefert.

Das Bundeskartellamt hat festgestellt, dass die Freitagspreise für die Dieselpreise am höchsten sind. Am Sonntag sinken die Preise, so dass sie sonntags und montags am niedrigsten sind (Bundeskartellamt, 2011). So lautet die erste Hypothese, die im Rahmen dieser Arbeit überprüft wird:

Hypothese 1: Der Freitagspreis der Kraftstoffpreise ist der höchste Preis der Woche.

Darüber hinaus hat das Bundeskartellamt festgestellt, dass die Spotpreise die Kraftstoffpreise mit einer Verzögerung beeinflussen (Bundeskartellamt, 2011). Insbesonde-

re schlagen die Rohölpreissteigerungen schneller auf die Kraftstoffpreise durch als die Rohölpreissenkungen (Wettbewerbskommission, 2008; Dudenhöfer, 2011).

Hypothese 2: Die Spotpreise beeinflussen die Benzinpreise mit einer Verzögerung.

Im zweiten Kapitel werden die modelltheoretischen Grundlagen und Methoden dargestellt. Es werden Stationarität und Saisonalität als wichtigste Charakteristika sowie die Grundlagen der Modellbildung der Zeitreihen vorgestellt.

Im dritten Kapitel wird die Analyse der empirischen Daten mittels Einheitswurzel-Tests auf Stationarität durchgeführt. Das Eliminieren der Nicht-Stationarität der Zeitreihen ist die Voraussetzung für eine gute Modellschätzung.

Im vierten Kapitel werden solche Charakteristika der Zeitreihen wie Periodizität und Saisonalität detailliert untersucht. Basierend auf den Ergebnissen dieser Untersuchung wird die Entscheidung zwischen dem additiven und multiplikativen Modell getroffen, um ein geeignetes Modell für die saisonbereinigte Zeitreihe zu generieren.

Im fünften Kapitel werden die Modelle gebildet, die die Entwicklung von Spotpreisen und Kraftstoffpreisen am Beispiel der Dieselpreise beschreiben. Der Zusammenhang zwischen den Spotpreisen und den Kraftstoffpreisen wird mittels eines Regressionsmodells geschätzt.

In der Zusammenfassung werden die Ergebnisse der Arbeit vorgestellt. Im Anhang finden sich die detaillierten Tabellen der Resultate der Tests auf Stationarität.

## 2 Analytische Grundlagen und Methoden

### 2.1 Zeitreihen und Zeitreihenmodelle

Die Daten, die der Ökonometrier in den Modellen analysiert, sind zum größten Teil Zeitreihen, also Folgen von Beobachtungen von Merkmalen, die in meist regelmäßigen Zeitpunkten erhoben werden. Nachdem die Beobachtungen in zeitlicher Ordnung erhoben werden, muss damit gerechnet werden, dass die jeweils aktuelle Beobachtung Information über die Vergangenheit enthält. Das aktuelle Konsumverhalten wird sich von dem in den letzten Perioden nicht allzu sehr unterscheiden. Ähnliches gilt für Investitionen oder andere ökonomische Variable. Wenn aber die Beobachtungen der jüngeren Vergangenheit zur Erklärung des aktuellen Verhaltens beitragen können, werden sie bei der Modellierung zu berücksichtigen sein, wenn von der in den Daten verfügbaren Information in effizienter Weise Gebrauch gemacht werden soll. Zeitreihen-Modelle, die von der Abhängigkeitsstruktur Gebrauch machen, haben sich insbesondere für das Prognostizieren von ökonomischen Variablen als sehr erfolgreich erwiesen (Hackl, 2005).

Die ARMA-Modelle einschließlich AR- und MA-Modelle beschreiben die aktuelle Beobachtung im Wesentlichen als gewichtete Summe verzögerter Realisationen der interessierenden Variablen. Zum begrifflichen Hintergrund von Zeitreihen-Modellen gehören die stochastischen Prozesse, ihre Eigenschaften wie die Stationarität und ihre Charakteristika wie die Autokorrelations- oder AC-Funktion und die Partielle Autokorrelations- oder PAC-Funktion sowie deren graphische Gegenstücke, das Korrelogramm und das Partielle Korrelogramm. Letztere eignen sich besonders zur Visualisierung der Abhängigkeitsstruktur der Variablen. Den Zeitreihen-Modellen wie den ARMA-Modellen entsprechen bestimmte Formen der Abhängigkeitsstruktur der Variablen, die durch die AC- und PAC-Funktionen charakterisiert werden können. Aus der Form dieser Funktionen kann auf den Modelltyp rückgeschlossen werden, durch den die Zeitreihe dargestellt werden kann.

Unter einer *Zeitreihe* versteht man eine Folge von Beobachtungen einer Zufallsvariablen  $Y$ , beispielsweise die jährlichen Werte des BIP, die quartalsweisen Zuwächse des Kapitalbestandes oder die monatlichen Importe. Hat man  $n$  solche Beobachtungen, so kann man sie als Realisationen der Zufallsvariablen  $Y_1, \dots, Y_n$  schreiben. Bei einer Zeitreihe spricht man auch von der Realisation eines stochastischen Prozesses; die der Zeitreihe entsprechenden Zufallsvariablen  $Y_1, \dots, Y_n$  sind ein Ausschnitt aus der unendlichen Folge  $\{Y_t, t = -\infty, \dots, \infty\}$  von Zufallsvariablen, die den stochastischen Prozess repräsentiert (Hackl, 2005).

Von einem Zeitreihen-Modell erwartet man, dass es in der Lage ist, die Charakteristika wie Trend, Saisonalität und irreguläre Fluktuationen darzustellen. Dazu stehen unterschiedliche Typen oder Klassen von Modellen zur Verfügung. Ein Modell kann darauf zielen, den Verlauf der Zeitreihe in globaler Weise zu beschreiben. Für die Entwicklung

des persönlich verfügbaren Einkommens  $Y$  könnte ein Modell lauten:

$$Y_t = \beta t + \sum_{i=1}^4 \gamma_i D_{it} + u_t, \quad (2.1)$$

wobei die  $i$ -te Dummy-Variable  $D_{it}$  den Wert Eins hat, wenn  $t$  dem Quartal  $i$  entspricht, und ansonsten den Wert Null hat. Andere Modelle beschreiben nur das lokale Verhalten und passen sich flexibel den sich ändernden Trends, Saisonmuster etc. an (Hackl, 2005).

## 2.2 Stationarität

Die Variablen  $Y_t$  entwickeln sich typischerweise in Abhängigkeit voneinander, wobei diese Abhängigkeit durch eine gemeinsame Wahrscheinlichkeitsverteilung  $p(y_1, \dots, y_n)$  beschrieben wird. Das wesentliche Charakteristikum der Verteilung  $p(\cdot)$  ist der Verlauf des Erwartungswertes  $\mu_t = E\{Y_t\}$ . Dieser Verlauf wird zum Extrapolieren einer Zeitreihe zum Zweck der Prognose verwendet. Für die Beurteilung eines Prozesses wichtigste Information enthalten

- die Varianzen  $Var\{Y_t\}$  und
- Kovarianzen  $Cov\{Y_t, Y_{t+k}\} = E\{(Y_t - \mu_t)(Y_{t+k} - \mu_{t+k})\}$ ,

die man für jedes Paar von Zeitpunkten  $t$  und  $t+k$  mit beliebigem  $k$  aus der Verteilung  $p(\cdot)$  ableiten kann. Wenn die Kovarianz-Funktion  $\gamma_t, k = Cov\{Y_t, Y_{t+k}\}$ ,  $k = 0, \pm 1$ , für alle  $t$  die gleiche ist, heißt die Abhängigkeitsstruktur der Variablen  $Y_t$  für jede Folge von Zeitpunkten  $\{t, \dots, t+k\}$  - für beliebiges  $t$  und  $k$  die gleiche ist, dann spricht man von *Stationarität*.

Ein stochastischer Prozess  $\{Y_t, t = -\infty, \dots, +\infty\}$  ist *stationär*, wenn:

- $E\{Y_t\} = \mu_t$
- $Cov\{Y_t, Y_{t+k}\} = \gamma_k$  für alle  $t$  und alle  $k$ . Die Kovarianz-Funktion des stationären Prozesses  $\gamma_k$  hängt also nur von  $k$  ab (Hackl, 2005).

### 2.2.1 Deterministische und stochastische Trends

Ökonomische Zeitreihen sind Realisationen von stochastischen Prozessen. Die meisten ökonomischen Zeitreihen zeigen einen Trend und sind die Realisation eines nicht-stationären Prozesses. Die Regressionsmodelle bilden den Zusammenhang zwischen den abhängigen Variablen und ihren erklärenden Regressoren als statische Beziehung ab. Das Vorhandensein der speziellen Abhängigkeitsstruktur und von Trends legt es aber nahe, die in diesen Spezifika der Daten enthaltene Information in der Modellierung zu

berücksichtigen. Entsprechende Modelle werden seit etwa 25 Jahren in großem Umfang eingesetzt und diskutiert.

Wenn die Nicht-Stationarität in einem Trend besteht, unterscheidet man zwischen trend-stationären und differenz-stationären Prozessen. Beide sind nicht-stationäre Prozesse, doch unterscheiden sich in der Art, wie die Nicht-Stationarität in der Modellierung berücksichtigt werden kann (Hackl, 2005).

Ökonomische Größen wie Einkommen, Konsum, Geldmengen etc. weisen in den meisten Wirtschaftsräumen ein stetes Wachstum auf. Demographische, technologische und andere Faktoren sind die Ursachen dafür. Die Verknappung der Erdölressourcen kann einen wichtigen Grund für das Wachstum der Benzinpreise sein. Das Darstellen von Trends ist eine zentrale Aufgabe der ökonometrischen Modellierung. Ein Fall von Nicht-Stationarität ist Nicht-Konstanz des Erwartungswertes. Wenn die Nicht-Konstanz bei Fortschreiten der Zeit zu immer größeren oder immer kleineren Erwartungswerten führt, so spricht man von einem Trend.

Ein *deterministischer Trend* eines Prozesses  $Y_t$  ist eine Funktion  $f_t$  der Zeit  $t$ , die den Erwartungswert von  $Y$  oder eine Komponente von  $Y$  beschreibt.  $Y_t$  kann man als  $Y_t=f(t)+u_t$  mit Weißem Rauschen  $u$  darstellen; die Varianz der  $u_t$  ist  $\sigma^2$ . Ein häufig verwendetes Modell für einen Prozess mit deterministischem Trend ist der lineare Trend:

$$Y_t=\alpha + \beta t + u_t \quad (2.2)$$

Die Koeffizienten  $\alpha$  und  $\beta$  geben den Wachstumspfad vor; der Trend kann ein steigender ( $\beta > 0$ ) oder ein fallender sein. Die Funktion  $f(t)$  kann ein quadratisches Polynom in  $t$  oder eine Exponentialfunktion oder eine andere Funktion von  $t$  sein (Hackl, 2005).

*Stochastischer Trend* im Modell

$$Y_t=\delta+Y_{t-1}+u_t \quad (2.3)$$

oder  $\Delta Y_t=Y_t-Y_{t-1}=\delta+u_t$  mit Weißem Rauschen  $u$  beschreibt ein irreguläres oder zufälliges Fluktuieren der Differenzen  $\Delta Y_t$  um den Erwartungswert  $\delta$ . Ein Prozess, der diesem Modell folgt, heißt *Random walk* mit Trend. Das Einsetzen der Modellgleichung für  $Y_{t-1}$ ,  $Y_{t-2}$  etc. liefert die Beziehung

$$Y_t=Y_0+\delta t+\sum_{i \leq t} u_i, \quad (2.4)$$

wobei  $\delta$  den Trendparameter darstellt. Der Prozess setzt sich aus zwei Komponenten zusammen, aus dem Wachstumspfad  $Y_0+\delta t$  und den kumulierten Störgrößen  $\sum_i u_i$  (Hackl, 2005).

Die Eigenschaften von *Random Walk* sind:

- für den Erwartungswert  $E\{Y_t\} = Y_0 + \delta t$  kein fixer Wert für alle  $t$
- für die Varianz der kumulierten Störgrößen  $\sum_{i \leq t} u_i$  bzw. für die Varianz von  $Y_t$  ergibt sich  $Var\{Y_t\} = \sigma^2 t$ ; die Varianz von  $Y_t$  wird mit wachsendem  $t$  beliebig groß
- Korrelation zwischen  $Y_t$  und  $Y_{t-k}$  ist  $\varrho_{t,k} = Corr\{Y_t, Y_{t-k}\} = \sqrt{1 - \frac{k}{t}}$

Für fixes  $k$  sind die  $Y_t$  und  $Y_{t-k}$  umso stärker korreliert, je größer  $t$  ist. Für wachsendes  $k$  strebt  $\varrho_{t,k}$  gegen den Wert Null. Allerdings ist die Geschwindigkeit dieser Konvergenz umso kleiner, je größer  $t$  wird. Man sagt, der Random Walk hat ein langes Gedächtnis. Nicht-Konstanz betrifft Erwartungswert, Varianz und Abhängigkeitsstruktur.

Das Eliminieren des Trends hat das Ziel,  $Y$  in einen stationären Prozess zu überführen. Im Fall von einem Random Walk,  $Y_t = \delta + Y_{t-1} + u_t$ , einem differenz-stationären Prozess, kann man einen stationären Prozess  $\Delta Y_t = \delta + u_t$  durch das Bilden von ersten Differenzen ableiten. Die Ordnung der Integration ist Eins, wenn die ersten Differenzen ein stationärer Prozess sind (Hackl, 2005).

### 2.2.2 Einheitswurzel-Tests

Wenn ein Trend in einem Modell korrekt berücksichtigt werden soll, muss bekannt sein, ob es sich um einen deterministischen oder einen stochastischen Trend handelt. Ein wichtiges Hilfsmittel dazu sind die *Einheitswurzel-Tests*, mit deren Hilfe das Vorliegen von nicht-stationären Zeitreihen diagnostiziert werden kann.

Man geht von einer Variablen  $Y$  aus, welche die Realisation eines Prozesses ist, der dem Modell  $Y_t = \varphi Y_{t-1} + u_t$  folgt;  $u$  ist Weißes Rauschen mit der Varianz  $\sigma^2$ . Ist  $Y$  ein differenz-stationärer Prozess, so hat  $\varphi$  den Wert Eins:  $Y$  ist dann ein Random Walk. Zum Prüfen, ob es zutrifft, dass  $Y$  ein Random Walk ist, testet man die Nullhypothese  $H_0: \varphi = 1$  gegen die Alternative  $H_1: \varphi < 1$ .

Der OLS-Schätzer von  $\varphi$  ergibt sich zu:  $\hat{\varphi} = \frac{\sum_{t=2}^n Y_t Y_{t-1}}{\sum_{t=2}^n Y_{t-1}^2}$ . Zum Testen von  $H_0$  bietet

sich die  $t$ -Statistik an:  $\tau = \frac{\hat{\varphi} - 1}{se(\hat{\varphi})}$ .  $se(\hat{\varphi})$  ist der Standardfehler des Schätzers  $\hat{\varphi}$ . Wenn  $|\varphi| < 1$ , ist  $Y_t$  stationär, so ist  $\tau$  die Teststatistik des  $t$ -Tests:  $\tau \sim t(n-1)$ . Ist  $Y_t$  nicht-stationär, so gilt  $\varphi=1$ , so folgt  $\tau$  auch bei Zutreffen der Nullhypothese nicht der  $t$ -Verteilung. Fuller (1976) gibt Perzentile für die Verteilung von  $\tau$  unter  $H_0$  an, die aus Monte Carlo Simulationen geschätzt worden sind. Auf der Basis von kritischen Schranken nach Fuller kann die Entscheidung zwischen  $H_0: \varphi=1$  und  $H_1: \varphi < 1$  mit der  $\tau$ -Statistik in Analogie zum  $t$ -Test getroffen werden. Der Name *Einheitswurzel* kommt daher, dass die Nullhypothese dem Fall entspricht, dass das Charakteristische Polynom



zum Modell  $Y_t = \varphi Y_{t-1} + u_t$ ,  $\Phi(z) = 1 - \varphi \cdot z$  die Wurzel  $z_1 = 1$  hat (Hackl, 2005).

Der Einheitswuzel-Test wurde erstmals von Dickey & Fuller (1979) diskutiert; daher wird der Test auch als Dickey-Fuller-Test bezeichnet. Das Modell

$$Y_t = \varphi Y_{t-1} + u_t \quad (2.5)$$

kann man auch schreiben als

$$\Delta Y_t = (\varphi - 1) Y_{t-1} + u_t = \delta Y_{t-1} + u_t. \quad (2.6)$$

Der Test von  $H_0: \delta=0$  gegen die Alternative  $H_1: \delta < 0$  ist äquivalent dem oben geschriebenen Einheitswuzel-Test von  $H_0: \varphi=1$  und  $H_1: \varphi < 1$ . Zum Entscheiden über  $H_0$  verwendet man die Teststatistik  $\tau = \frac{d}{se(d)}$ .

Der OLS-Schätzer  $d$  für  $\delta$  ergibt sich zu  $d = \hat{\varphi} - 1$ , wobei  $\hat{\varphi}$  der OLS-Schätzer von  $\varphi$  ist. Für den Standardfehler  $se(d)$  gilt  $se(d) = se(\hat{\varphi} - 1) = se(\hat{\varphi})$  (Hackl, 2005).

Nicht alle Zeitreihenvariablen kann man als einen AR-Prozess 1. Ordnung darstellen. Es gibt eine Möglichkeit, den Dickey-Fuller-Test für die Prozesse höherer Ordnung anzuwenden, so ist der *erweiterte Dickey-Fuller-Test*. Ausgangspunkt des erweiterten Dickey-Fuller-Tests sind der AR(1)-Prozess

$$Y_t = \varphi Y_{t-1} + u_t \quad (2.7)$$

und die Einheitswuzel-Hypothese. Die Erweiterung ist das Testen der Unit-Root Hypothese für einen AR(p)-Prozess

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + u_t. \quad (2.8)$$

Diese Erweiterung berücksichtigt eine komplexere Abhängigkeitsstruktur der  $Y$ :

$$\Delta Y_t = \delta Y_{t-1} + \beta_1 \Delta Y_{t-1} + \dots + \beta_p \Delta Y_{t-p} + u_t \quad (2.9)$$

Die Koeffizienten sind Funktionen der  $\varphi_1$ ; z.B.  $\delta = \varphi_1 + \dots + \varphi_p - 1$ . Der Test von  $H_0: \delta=0$  wird wie der Dickey-Fuller-Test ausgeführt. Die Nullhypothese sagt die Nicht-Stationarität der Zeitreihe aus. Die Teststatistik  $\tau$  wird mit den gleichen kritischen Schranken verglichen. Das Modell kann um ein Interzept  $\alpha$  oder um einen deterministischen Trend  $\alpha + \beta t$  erweitert werden.

Die Ordnung  $p$  des AR(p)-Prozesses, den man dem Modell des ADF-Tests zugrunde legt, ist im Allgemeinen nicht bekannt. Man geht zum Auffinden des korrekten Modells vom allgemeinen zum engeren Modell vor. Dazu wird eine Obergrenze  $p_{max}$  der Ordnung  $p$  festgelegt. Eine Formel zum Bestimmen von  $p_{max}$ , die (Schwert, 1989) auf der Basis umfangreicher Monte Carlo Simulationen vorgeschlagen hat, ist (Hackl, 2005)

$$p_{max} = \text{int}\{12(n/100)^{0.25}\} \quad (2.10)$$

Beim Dickey-Fuller-Test nimmt man an, dass die Störgrößen unabhängig sind und über eine konstante Varianz verfügen. Im diesem Zusammenhang treten einige Probleme auf:

1. der Schätzer für  $\gamma$  und Standardfehler ist ungenau, bis alle AR-Prozesse in der Schätzungsgleichung berücksichtigt sind. Anders gesagt, wenn man die genaue Ordnung des AR-Prozesses nicht kennt, tritt das Problem der richtigen Wahl der Anzahl von  $p$ .
2. Der Dickey-Fuller-Test berücksichtigt nur eine Einheitswurzel. Das heißt, ein Prozess mit  $p$  Ordnungen und  $d < p$  Einheitswurzeln muss  $d$ -mal differenziert werden, um zur Stationarität zu gelangen.
3. Es gibt Zeitreihen, die erste Differenzierung und auch solche, die eine saisonale Differenzierung voraussetzen. Man muss eine Methode entwickeln, die zwischen den beiden Arten von Einheitswurzelprozessen unterscheiden kann.
4. Es können strukturelle Brüche in den Daten vorkommen, die einen scheinbaren Trend an die Daten vermitteln (Enders, 2010).

### 2.2.3 Residuen-Analyse

Unter bestimmten Annahmen über die Störungen  $U_t$  haben die geschätzten Koeffizienten  $\hat{\beta}_j$  Optimalitätseigenschaften. Sofern die  $U_t$  unabhängig normalverteilt mit Erwartungswert 0 und Varianz  $\sigma^2$  sind, handelt es sich bei Schätzern, die mit dem Kleinstquadratprinzip gewonnen wurden, um beste unverzerrte Schätzer. Werden diese Annahmen allerdings verletzt, so besitzen die Kleinstquadratschätzer keine oder nur noch unrealistisch eingeschränkte Optimalitätseigenschaften (wie „beste lineare Unverzerrtheit“) (Schlittgen & Streitberg, 1989). Es ist daher wichtig, die Gültigkeit der Annahmen jeweils kritisch zu überprüfen.

Grundlegend dabei sind die sogenannten *Residuen*, d.h. die Abweichungen von Daten  $x_t$  und geschätztem Trend  $\hat{m}_t$ . Unter Normalverteilungsannahmen sollte das Histogramm der Residuen näherungsweise einer Normalverteilung folgen - insbesondere sollten keine Ausreißer auftreten. Kleinstquadrateschätzer werden von untypischen Werten sehr stark beeinflusst. In eine quadratische Zielfunktion geht z.B. eine Abweichung 10 ebenso stark ein wie 100 Abweichungen der Größe 1. Gegebenenfalls ist es notwendig, vor der Schätzung von  $m_t$  etwaige Ausreißer zu beseitigen.

Der Erwartungswert aller  $U_t$  sollte gleich 0 sein - dies drückt inhaltlich aus, dass keine systematischen Effekte im Modell für  $m(t)$  vernachlässigt wurden. Überprüft werden kann diese Annahme dadurch, dass die Residuen  $\hat{u}_t$  gegen andere relevante Variablen, z.B. die Zeit  $t$  aufgetragen werden, um systematische Zusammenhänge zu entdecken. In Reihen mit Saisonschwankungen ist die Annahme  $E[U_t]=0$  bei einer Trendbestimmung im allgemeinen nicht erfüllt, da die Residuen  $\hat{u}_t = x_t - \hat{m}_t$  systematisch mit der Saison variieren. Es werden sich daher im Regelfall andere Trendschätzungen ergeben,

wenn man von vornherein ein gemeinsames Modell für Trend und Saison unterstellt. Man achte bei einer Zeichnung der  $\hat{u}_t$  gegen  $t$  insbesondere auf sogenannte Strukturbrüche, d.h. tiefgreifende Veränderungen des Gesamtbildes (Schlittgen & Streitberg, 1989).

Die Unabhängigkeitsannahme impliziert insbesondere, dass die Störungen  $u_t$  nicht autokorreliert sind. Die Residuen sollten zufällig um 0 schwanken. Allzu langsame Schwankungen deuten auf positive, allzu schnelle und regelmäßige Schwankungen auf negative Autokorrelation hin.

Ein Test auf Autokorrelation zum Lag (Verzögerung) 1 ist der *Durbin-Watson-Test*. Die Teststatistik ist

$$d = \sum_{t=2}^N (\hat{u}_t - \hat{u}_{t-1})^2 / \sum_{t=1}^N \hat{u}_t^2 \quad (2.11)$$

Die asymptotische Verteilung von  $d$  unter der Nullhypothese unkorrelierter Störungen ist bei Kendall (1971) angegeben. Werte von  $d$ , die nahe bei 2 liegen, deuten auf Unkorreliertheit hin, solche nahe bei 0 auf positive und Werte nahe 4 auf negative Korrelation. Insbesondere die Annahme der Unabhängigkeit ist in der Zeitreihenanalyse praktisch niemals erfüllt. Dies hat zu einem gewissen Misstrauen gegenüber Komponentenmodellen geführt und wesentlich die Hinwendung zur Theorie schwach stationärer Prozesse motiviert (Schlittgen & Streitberg, 1989).

## 2.3 Saisonalität

Saisonale Einflüsse gehören zu den markantesten Kennzeichen wirtschaftsstatistischer Zeitreihen. Vor allem quantitative Variablen (Einkommen, Konsum, Produktionsleistung, Lagerhaltung, Staatsausgaben, Beschäftigung, Geldmenge etc.) weisen ein zyklisches Verhalten innerhalb einzelner Jahre auf. Barsky & Miron (1989) zeigen anhand von Quartalsdaten für die USA im Zeitraum 1948-1985, dass die Standardabweichungen geschätzter saisonaler Dummy-Koeffizienten ( $\hat{\mu}_s$ ) einer einfachen Regression nach der Methode der Kleinsten-Quadrate (KQ-Schätzung)

$$x_t = \sum_{s=1}^S D_{s(t)} \mu_s + \varrho_t \quad (2.12)$$

Großteile der Streuung quantitativer Variablen  $x_t$  erklären (beispielsweise beim Konsum nichtdauerhafter Güter erklärten deterministische Saisonfiguren rund 97 Prozent der Variablenstreuung). Dabei waren  $x_t$  erste Differenzen logarithmierter Daten und  $\varrho_t$  ein als stationär unterstellter Residualprozess.  $D_s(t)$  repräsentiert eine sogenannte *saisonale Dummy-Variable* für eine Teilperiode  $s$ , die den Wert eins annimmt, wenn  $t$  zu eben dieser Teilperiode gehört und ansonsten gleich Null ist. Für Preisvariablen wie insbesondere Zinsen und Löhne spielten deterministische Saisoneinflüsse kaum eine Rolle bei der Varianzzerlegung. Bei Nominal- und Realzinsen betrugen die Streuungsanteile, z.B. nur 1 bzw. 6 Prozent (Herwatz, 1995).

Der Verwendung saisonaler Dummy-Variablen liegt die Vorstellung einer deterministischen Saisonfigur zugrunde, die strukturkonstant im Periodenverlauf beibehalten wird. Ein sich allmählich veränderndes saisonales Aktivitätsmuster kann so nicht repräsentiert werden. Schätzt man die obige Gleichung sukzessive für steigende Stichprobenumfänge  $T^*=S, S+1, S+2, \dots, T$  (rekursive KQ-Schätzung), so zeigen die Schätzkoeffizienten  $\hat{\mu}_s$  für bestimmte Variablen eine deutliche Bewegung im Zeitablauf. Streuungszerlegungen in der Art von Barsky & Miron (1989) bleiben damit rein deskriptiv.

Mögliche Ursachen für ein saisonal geprägtes Verlaufsmuster von Zeitreihen sind vielfältig, selten isoliert wirksam und zumeist einer wirtschaftspolitischen Steuerung entzogen. Sie reichen von Kalendereffekten (z.B. Weihnachtsfeiertage) und sogenannten „timing decisions“ (Ferienzeiten, Termine von Gewinnausschüttungen etc.) über klimatische Einflüsse bis hin zu Erwartungen über saisonales Verhalten der Wirtschaftssubjekte. Man beachte, dass sich einige dieser Verursacher im Zeitablauf verändern können, a priori spricht einiges für ein im Zeitablauf variables Bild saisonaler Strukturen.

In der spektralen Darstellung der Zeitreihe entspricht das beschriebene zyklische Verlaufsmuster einem hohen Anteil spektraler Dichte im Bereich der saisonalen Frequenzen. Derartige Spitzen im Zeitreihenspektrum definieren eine gegebene Zeitreihe als saisonal.

Für die Erklärung ökonomischer Zusammenhänge wird saisonalen Schwankungen zumeist wenig Bedeutung beigemessen. Saisonale Einflüsse gelten oft als „Variablenfehler“, zu deren Identifikation angenommen wird, dass sie ausschließlich zur spektralen Masse an saisonalen Frequenzen (bzw. zu einem engen umliegenden Frequenzband) beitragen. Implizit geht man also davon aus, dass Saisonzyklen einerseits und längerfristige Zyklen andererseits unabhängig voneinander sind (Orthogonalitätsannahme). Begründet liegt diese Minderschätzung saisonaler Verlaufsmuster sicher mit in dem hohen Interesse der quantitativen Forschung an Wachstums- und Konjunkturmodellen. Vor diesem Hintergrund ist es verständlich, dass vergleichsweise gut prognostizierbare Niveauschwankungen innerhalb einer Periode geringere Bedeutung erhalten als eine als dauerhaft unterstellte Veränderung aggregierter wirtschaftlicher Aktivität (Herwatz, 1995).

### 2.3.1 Saisonbereinigung von Zeitreihen

Vielen traditionellen statistischen Modellen zufolge zerfallen ökonomische Zeitreihen in zwei unkorrelierte Komponenten, eine saisonale und eine nicht saisonale. Für beide Komponenten werden unterschiedliche datengenerierende Modelle angenommen. Entsprechend verbreitet ist die saisonale Bereinigung ökonomischer Zeitreihen.

Die Modellierung saisonal bereinigter Daten kann den Vorteil haben, dass die Anpassung des Modells an die Daten nicht dominiert wird von der Konzentration spektraler Dichte an saisonalen zuungunsten anderer Frequenzen. Dies ist insbesondere bei sol-

chen Schätzungen nicht der Fall, die den Daten ein restringiertes Modell anpassen. Sims (1974) zeigt mit Hilfe der Fouriertransformation der Modellkoeffizienten im sogenannten „distributed-lag“-Modell für zwei Zeitreihen ökonomischer Variablen  $w_t$ ,  $x_t$ ,

$$w_t = \mu + x_t \beta_0 + x_{t-1} \beta_1 + \dots + u_t, \quad (2.13)$$

das auf beobachtbare, von saisonalen Einflüssen geprägte Daten angewendet wird, die Größenordnung der Abweichung von „wahren“ d.h. hier saisonfreien Wirkungszusammenhang. Die Tatsache, dass saisonale Komponenten die spektrale Dichte annahmegoemäß nur an saisonalen Frequenzen beeinträchtigen, hat zur Folge, dass die Verzerrung der Schätzkoeffizienten nur an saisonalen Frequenzen von Bedeutung ist. Besonders schwerwiegend kann diese Verzerrung werden, falls den Schätzkoeffizienten Restriktionen auferlegt werden. Unter bestimmten Bedingungen werden saisonal bedingte Verzerrungen abgemildert:

1. Der strukturelle Zusammenhang ist für die interessierenden d.h. saisonbereinigten Variablen der gleiche wie für die saisonalen Komponenten. Eine Verzerrung tritt nicht auf.
2. Die Saisonkomponenten der abhängigen Variablen einerseits und der unabhängigen Variablen andererseits korrelieren nicht oder kaum miteinander (Herwatz, 1995).

Der Bereinigung saisonaler Daten liegt die Annahme zugrunde, dass saisonale Datenvariationen für eine Reihe von Fragestellungen der empirischen oder auch theoretischen Arbeit von untergeordneter Bedeutung sind. Die Bereinigung saisonaler Einflüsse ist jedoch z.B. von folgenden Problemen belastet:

1. Selbst wenn saisonale Einflüsse als Variablenfehler einzelner Zeitreihen behandelt werden können, so ist dennoch die Annahme kritisch, dass diese Fehler unabhängig voneinander in verschiedenen Zeitreihen auftreten. Bei „abhängigen Fehlern“ beraubt sich aber die quantitative Analyse im multiplen Modell von vornherein bereits möglicher Erkenntnisse struktureller Zusammenhänge.
2. Zeitreihenfilter zur Saisonbereinigung beziehen in aller Regel für jeden Zeitpunkt  $t$ , sowohl vergangene als auch zukünftige Realisierungen von  $x_t$  ein (z.B. gleitende Durchschnitte). Zumeist sind derartige Filter symmetrisch, d.h. die Beobachtungen  $x_{t+h}$  und  $x_{t-h}$ ,  $h=0,1,2,\dots$  erhalten bei der Kalkulation des saisonbereinigten Wertes  $x^*_t$  jeweils gleiche Gewichte. Die Verwendung eines solchen zweiseitigen linearen Filters zum Erhalt der saisonbereinigten Reihe  $x^*_t = a * (B)x_t$  ist am aktuellen Rand der Zeitreihe nicht möglich. Einseitige Filter dienen in diesem Bereich zur Ermittlung vorläufiger Werte, die gegebenenfalls dann später korrigiert werden müssen (Herwatz, 1995).

### 2.3.2 Modelle für saisonale Daten

Die Box-Jenkins Technik für die Modellierung der saisonalen Daten ist etwas andere als die für die nicht-saisonalen Daten. Die Saisonalität in den Daten mit der Periode  $s$  entsteht eher durch die saisonalen Koeffizienten der AC-Funktion und der PAC-Funktion der zeitlichen Verzögerungen  $s, 2s, 3s, \dots$ , als die der  $1, 2, 3$ . Zwei reine saisonale Modelle der Quartalsdaten sind in den Gleichungen 2.14 und 2.15 dargestellt.

$$y_t = a_4 y_{t-4} + \varepsilon_t, \|a_4\| < 1 \text{ und} \quad (2.14)$$

$$y_t = \varepsilon_t + \beta_4 \varepsilon_{t-4} \quad (2.15)$$

Das theoretische Korrelogramm für die erste Gleichung zeigt, dass  $p = (a_4)^{i/4}$ , wenn  $i/4$  ein Integer ist und  $p_i = 0$ , wenn sonst. Die AC-Funktion weist eine Spitze an den Verzögerungsstellen  $4, 8, 12, \dots$  auf. In der Praxis wird die Identifikation durch die Tatsache erschwert, dass die saisonale Komponente mit der nicht-saisonalen Komponente in den Daten in Verbindung gerät. Die AC- und PAC-Funktionen widerspiegeln zugleich beide Komponenten. Für Quartalsdaten, der saisonale MA-Term wird folgende Form haben:

$$y_t = a_1 y_{t-1} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_4 \varepsilon_{t-4} \quad (2.16)$$

Eine Alternative wäre, von dem autoregressiven Koeffizienten der 4. Verzögerungsstelle Gebrauch zu machen, um die Saisonalität zu erfassen:

$$y_t = a_1 y_{t-1} + a_4 y_{t-4} + \varepsilon_t + \beta_1 \varepsilon_{t-1} \quad (2.17)$$

Beide Methoden in den Gleichungen 2.16 und 2.17 sind additive Methoden der saisonalen Koeffizienten: ein AR- oder MA-Koeffizient wurde in der Saisonalität berücksichtigt. Die multiplikative Methode ermöglicht das Zusammenwirken vom ARMA-Modell und der saisonalen Effekte, siehe Gleichung 2.18 (Enders, 2010).

$$(1 - a_1 \cdot L)y_t = (1 + \beta_1 \cdot L)(1 + \beta_4 \cdot L^4) \varepsilon_t \quad (2.18)$$

## 2.4 Transformation von Zeitreihen durch Filter

### 2.4.1 Methode der gleitenden Durchschnitte

Neben der Bestimmung des globalen Trends ist oft eine Glättung der Zeitreihe von Interesse. Glättung bedeutet Ausschaltung von irregulären Schwankungen durch lokale Approximationen. Der Vorteil liegt u.a. darin, dass man mit Polynomen niedrigen Grades auskommt. Im klassischen Komponentenmodell entspricht dies der näherungsweise Bestimmung der glatten Komponente. Es ist naheliegend, im ersten Ansatz zum Glätten einer Zeitreihe die Beobachtung  $x_t$  durch ein lokales arithmetisches Mittel  $y_t$  zu ersetzen:

$$y_t = \frac{1}{2q+1} \sum_{u=-q}^q x_{t-u}, \quad t = q+1, \dots, N-q \quad (2.19)$$

Dies ist ein Beispiel für eine lineare Transformation einer Zeitreihe  $(x_t)$  in eine andere  $(y_t)$ . Eine lineare Transformation  $L$  einer Zeitreihe  $(x_t)$  in eine andere  $(y_t)$  gemäß

$$y_t = L \cdot x_t = \sum_{u=-q}^s a_u \cdot x_{t-u}, \quad t = s+1, \dots, N-q \quad (2.20)$$

wird als linearer Filter  $L$  bezeichnet. Anstatt durch das formale Symbol  $L$  wird der Filter auch durch seine Gewichte  $a_u$  in der Form  $(a_{-q}, \dots, a_s)$  bzw.  $(a_u)$  angegeben. Die Anwendung eines Filters auf eine Zeitreihe  $(x_t)$  wird als *Filtration der Reihe* bezeichnet.  $(x_t)$  heißt auch der Input und die gefilterte Reihe  $(y_t)$  der Output des Filters. Bei der Filtration einer Zeitreihe  $(x_t)$  ist zu beachten, dass die gefilterte Reihe  $(y_t)$  kürzer als die Input-Reihe ist. Im Fall  $s > 0$  wird der Anfang, im Fall  $q > 0$  das Ende gekappt (Schlittgen & Streitberg, 1989).

Dieses Verfahren zur Glättung der Zeitreihen wird als *einfacher gleitender Durchschnitt* bezeichnet. Der entsprechende Filter hat die Gewichte  $a_u = 1/(2q+1)$  mit  $\sum a_u = 1$ . Ein linearer Filter  $(a_u)$  mit  $\sum a_u = 1$  heißt ein *gleitender Durchschnitt*. Im Fall  $a_u = 1/(2q+1)$ ,  $u = -q, \dots, q$ , spricht man von einem *einfachen gleitenden Durchschnitt*. Beispiel von zwei einfachen gleitenden Durchschnitten, bei denen je drei bzw. fünf Werte einbezogen wurden, sieht folgenderweise aus:

- $y_t = \frac{x_{t-1} + x_t + x_{t+1}}{3}$
- $z_t = \frac{x_{t-2} + x_{t-1} + x_t + x_{t+1} + x_{t+2}}{5}$

Einfache gleitende Durchschnitte können auch für gerade Anzahlen von Werten bestimmt werden. Der Output ist dann aber jeweils der Mitte zwischen zwei Zeitpunkten zugeordnet, so etwa  $1/4(x_1 + x_2 + x_3)$  dem Zeitpunkt 2.5. Man kann diesen ungewünschten Effekt dadurch vermeiden, dass man jeweils über zwei aufeinander folgende Werte des Outputs mittelt:

- $y_{2.5} = \frac{x_1 + x_2 + x_3 + x_4}{4}$  und
- $y_{3.5} = \frac{x_2 + x_3 + x_4 + x_5}{4}$ , als Mittelwert entsteht
- $y_3 = 1/2 y_{2.5} + 1/2 y_{3.5} = 1/8 x_1 + 1/4 x_2 + 1/4 x_3 + 1/4 x_4 + 1/8 x_5$ .

Die lokale Approximation einer Zeitreihe durch Polynome führt ebenfalls zu gleitenden Durchschnitten, wenn jeweils nur der für den mittleren Zeitpunkt erhaltene Wert  $y_t$  des Polynoms als geglätteter Wert anstelle von  $x_t$  genommen wird (Schlittgen & Streitberg, 1989).

### 2.4.2 Methode des Differenzenfilters

Bei der Anpassung eines Polynoms an eine Zeitreihe und bei der lokalen Glättung durch Polynome stellte sich die Frage, welcher Grad für das Polynom gewählt werden soll.  $f(t)$  sei ein Polynom vom Grade  $p > 0$ :  $f(t) = a_0 + a_1t + \dots + a_pt^p$ . Dann ist  $g(t) = f(t) - f(t-1)$  ein Polynom vom Grade höchstens  $p-1$ . Die Differenzenbildung reduziert den Polynomgrad um 1. Wird das Verfahren  $p-1$ -mal angewendet, so erhält man bei einem Polynom  $p$ -ten Grades einen konstanten Wert. Die Anwendung der Differenzenbildung auf Zeitreihen führt wieder auf eine lineare Transformation oder Filtration.

Der lineare Filter  $\Delta$ , der definiert ist durch  $\Delta x_t = x_t - x_{t-1}$ ,  $t = 2, 3, \dots, N$  heißt *Differenzenfilter 1. Ordnung*. Differenzenfilter  $p$ -ter Ordnung,  $p > 1$ , sind rekursiv definiert durch  $\Delta^p x_t = \Delta^{p-1} x_t - \Delta^{p-1} x_{t-1}$ ,  $t = p+1, \dots, N$ . Ist bei dem der Trendbestimmung zugrundegelegten Ansatz ein Polynom, so lässt sich durch Differenzenbildung der Trend eliminieren bzw. der Grad des Polynoms bestimmen.

Für ökonomische Zeitreihen ist es oft ausreichend, die Differenzen 1. Ordnung zu bilden, um Veränderungen im Niveau zu beseitigen. Da die ersten Differenzen als Zuwächse auch inhaltlich bedeutsam sind, wird dieses Vorgehen häufig praktiziert. Falls nötig, werden die zweiten Differenzen gebildet, um zusätzliche Nicht-Stationarität der Steigungen zu beseitigen (Schlittgen & Streitberg, 1989).

## 2.5 Modellschätzung

### 2.5.1 ARMA-Modelle

Eine sehr allgemeine Klasse von Modellen sind die *ARMA-Modelle*. Die ARMA-Modelle sind eine Klasse von Zeitreihen-Modellen, die beispielsweise für die Prognose eine wichtige Rolle spielen. Ihre Bedeutung verdanken sie auch der von Box und Jenkins (1976) entwickelten Technik, die in der Literatur umfassend dokumentiert ist und deren Anwendung durch die Verfügbarkeit von exzellenten Software-Paketen unterstützt wird. Für die Zwecke der trigonometrischen Modellierung sind die Möglichkeiten der ARMA-Modelle als Analyseverfahren weniger interessant als die mit den Modellen verbundenen AC- und PAC-Funktionen, die ein wichtiges Instrument zum Spezifizieren eines geeigneten Modells ist (Hackl, 2005).

Es ist möglich zum ARMA-Modell zu gelangen, indem man den Moving-Average Prozess mit der linearen Differentialgleichung zusammenführt. Wenn man die  $p$ -Ordnung der Differentialgleichung berücksichtigt, so sieht das Modell wie in der Gleichung 2.21 aus.

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + x_t, \quad (2.21)$$



Und wenn  $x_t$  ein Moving Average Prozess von der Ordnung  $q$  ist, wird das Modell folgende Form haben:

$$y_t = a_0 + \sum_{i=1}^p a_i y_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i} \quad (2.22)$$

Wenn sich die charakteristischen Wurzeln der zweiten Gleichung im Einheitskreis befinden, liegt ein ARMA-Modell für  $y_t$  vor. Der autoregressive Teil des Modells ist die Differentialgleichung, gegeben durch den homogenen Teil der ersten Gleichung. Wenn er die  $p$ -Ordnung enthält und das Modell für  $x_t$  der  $q$ -Ordnung ist, heißt das Modell das ARMA (p,q)-Modell. Wenn  $q=0$  ist, so ist es ein reiner *autoregressiver Prozess*  $AR(p)$ , und wenn  $p=0$ , liegt ein reiner *moving-average Prozess*  $MA(q)$ . Wenn sich die charakteristischen Wurzel der zweiten Gleichung nicht im Einheitskreis befinden, so ist  $y_t$  ein *autoregressiver integrierter moving-average (ARIMA) Modell* (Enders, 2010).

Wie aus einem ARMA-Modell ein AR-Modell und MA-Modell hergeleitet werden kann, wird im folgenden dargestellt. Die allgemeine Form eines ARMA(p,q)-Modells lautet:

$$Y_t = \alpha + \gamma_1 Y_{t-1} + \dots + \gamma_p Y_{t-p} + \varepsilon_t, \varepsilon_t = u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}, \quad (2.23)$$

wobei die Variablen  $u$  Weißes Rauschen sind, also identisch und unabhängig verteilte Variable mit Erwartungswert Null und Varianz  $\sigma^2$ .

Voraussetzung für die Darstellbarkeit des ARMA(p,q)-Prozesses als  $MA(\infty)$ -Prozess ist die Stationarität des AR-Teils  $Y_t = Y_{t-1} + \varepsilon_t$ .  $u$  wird als gewichtete Summe von verzögerten Komponenten des Weißen Rauschens dargestellt, indem man wiederholt die Modellgleichung verwendet, um verzögerte  $u$ 's zu eliminieren:

$$Y_t = \psi_0 + \sum_{i=1}^{\infty} \psi_i u_{t-i} \quad (2.24)$$

$\sum_{i=1}^{\infty} \psi_i^2 < \infty$  als notwendige Bedingung für die Stationarität von  $Y$ , erfüllt wenn für die Wurzeln  $z_i$ ,  $i = 1, \dots, p$  des Charakteristischen Polynoms  $\Phi(z) = 1 - \gamma_1 z - \dots - \gamma_p z^p$  gilt  $|z_i| > 1$ . Die Koeffizienten  $\psi$  sind Funktionen der Parameter  $\gamma_1, \dots, \gamma_p$  und  $\theta_1, \dots, \theta_q$ .

Voraussetzung für die Darstellbarkeit des ARMA(p,q)-Prozesses als  $AR(\infty)$ -Prozess ist die Invertierbarkeit des MA-Teils  $Y_t = \alpha + \sum_{i=1}^{\infty} \varepsilon_{t-i}$ . Dann lässt sich das  $AR(\infty)$ -Modell als unendliche Summe verzögerter  $Y_{t-1}$  darstellen. Diese Umformung setzt voraus, dass für die Wurzeln  $z_i$ ,  $i = 1, \dots, p$  des Charakteristischen Polynoms  $\Theta(z) = 1 - \theta_1 z - \dots - \theta_p z^p$  gilt  $|z_i| > 1$ . Diese Bedingung heißt *Invertierbarkeits-Bedingung* (Hackl, 2005).

### 2.5.2 AR(p)-Prozess

Folgt  $Y$  einem AR(p)-Prozess, so gilt für  $Y_t$  die Beziehung

$$Y_t = \alpha + \gamma_1 Y_{t-1} + \dots + \gamma_p Y_{t-p} + u_t, u_t \sim \text{IID}(0, \sigma^2) \quad (2.25)$$

Die Variable  $Y$  wird also auf ihre eigenen verzögerten Werte regressiert. Für den AR(p)-Prozess ist die Invertierbarkeit immer gegeben. Für den AR(1)-Prozess müssen bestimmte Stationaritäts-Bedingungen erfüllt sein: Für die Wurzeln  $z_i$  des Charakteristischen Polynoms  $\Phi(z) = 1 - \gamma_1 z - \dots - \gamma_p z^p$  muss gelten:  $|z_i| > 1, i = 1, \dots, p$ .

Die AC-Funktion des AR(p)-Prozesses verläuft exponentiell abnehmend mit oder ohne alternierende Werte oder in Form einer gedämpften Sinusschwingung. Die PAC-Funktion ist nur für  $k \leq p$  von Null verschieden. Ab  $k = p + 1$  hat sie den Wert Null; die PAC-Funktion bricht also in  $p$  ab (Hackl, 2005).

### 2.5.3 MA-Prozess

Der Name MA-Modell kommt von *moving average*, also gleitende Mittelung, jene Operation, welche die Generierung der  $Y$  charakterisiert. Unter dem MA(q)-Modell versteht man ein ARMA(p,q)-Modell mit  $p=0$ , also ein Modell, das die Variable als gewichtete Summe von verzögerten Komponenten des Weißen Rauschens beschreibt. Der entsprechende MA-Prozess ist eine Folge von Zufallsvariablen  $Y_t, t=1, \dots, n$ , die entsprechend dem MA-Modell generiert werden.

Der MA(1)-Prozess ist eine Folge von Zufallsvariablen  $Y_t$ , für die gilt

$$Y_t = \alpha + u_t - \theta u_{t-1}, u_t \sim \text{IID}(0, \sigma^2) \quad (2.26)$$

Die Verallgemeinerung des MA(1)-Prozesses ergibt den MA(q)-Prozess

$$Y_t = \alpha + u_t - \theta u_{t-1} - \dots - \theta_q u_{t-q}, u_t \sim \text{IID}(0, \sigma^2). \quad (2.27)$$

Er hat folgende Eigenschaften:

- Der MA(q)-Prozess ist stationär.
- Seine AC-Funktion bricht mit  $q$  ab:  $\rho_k = 0$  für  $k > q$ .

Für die praktische Anwendung ist das Abbrechen der AC-Funktion mit  $q$  ein wichtiges Merkmal (Hackl, 2005).

## 3 Analyse der Zeitreihen auf Stationarität

### 3.1 Vorbereitung der Daten

Die Daten zu den Diesel-, Benzin-, und SuperPluspreisen wurden im Zeitraum von 31.05.2009 bis 20.08.2010 von einer freien Tankstelle Nitrol in Lengerich im Bundesland Nordrhein-Westfalen erhoben. Das Ablesen der Daten hat jeweils täglich um 18:00 erfolgt.

Der Datensatz zu Spotpreisen des Rohstoffes Erdöl (Dollar per Barrel) wurde für den gleichen Zeitraum von der Datenbank auf der Holding Webseite (Holding Company „Finam“) heruntergeladen. Es wurde festgestellt, dass die Datenbanken keine Informationen über die Samstagpreise liefern. Im Zeitraum von 31.05.2009 bis 20.08.2010 stehen insgesamt 384 Beobachtungen zur Verfügung. Zunächst wurde die Währungsumrechnung Dollar in Euro durchgeführt, um die den Einfluss der Wechselkursschwankungen im Modell auszuschließen. Die Wechselkursdaten sind der gleichen Datenbank entnommen. Die Maßumrechnung Barrel in Liter wurde mit  $1 \text{ Barrel} = 158,98722 \text{ Liter}$  vollendet.

Den Datensatz zu den Diesel-, Benzin-, und SuperPluspreisen wurde an den Datensatz der Spotpreise durch das Eliminieren von Samstagpreisen angepasst und um die Steuerkomponente bereinigt. Im Zeitraum der Datenerhebung hat keine Änderung der Energiesteuersätze erfolgt, somit ist die Steuerkomponente konstant und kann eliminiert werden. Der Benzinpreis am 23. Mai 2009 um 18:00 beträgt 1,299 Euro/Liter. Durch das Eliminieren der Mehrwertsteuer in der Höhe von 19 Prozent vom Produktpreis und der Mineralölsteuer mit Ökosteuern, die konstant auf jeden Liter des bereinigten um Mehrwertsteuer Produktpreises erhoben werden, kommt man rund auf 0,4371 Euro/Liter. Somit ergibt sich der Steueranteil von 0,636 Prozent. Aus dem Beispiel kann man gut schließen, dass die Steuerpolitik eine der wichtigsten Rollen in der Benzinpreiszusammensetzung spielt (Energieportal).

### 3.2 Vorläufige Analyse

Am Anfang ist die graphische Darstellung der Diesel-, Benzin-, Superplus- und Spotpreisen zu analysieren. In der Abbildung 1 sind die Originaldaten der Benzin-, Diesel-, SuperPluspreise per Liter dargestellt, die um die konstante Steuerkomponente bereinigt sind, und die Spotpreise in Euro auch per Liter. Man sieht deutlich an der Graphik, dass alle drei Kraftstoffpreise den Verlauf voneinander wiederholen. Es ist fast immer der gleiche Prozentsatz, um welchen sie sich voneinander unterscheiden. Man beobachtet auch einen ähnlichen Verlauf der Kraftstoffpreise und der Spotpreise. Es ist einen positiven Trend zu erkennen, der sich zum Ende des Zeitraums abschwächt. Im nächsten Schritt werden die Zeitreihen auf die Stationarität untersucht, um mögliche Trends oder Saisonalitäten festzustellen.

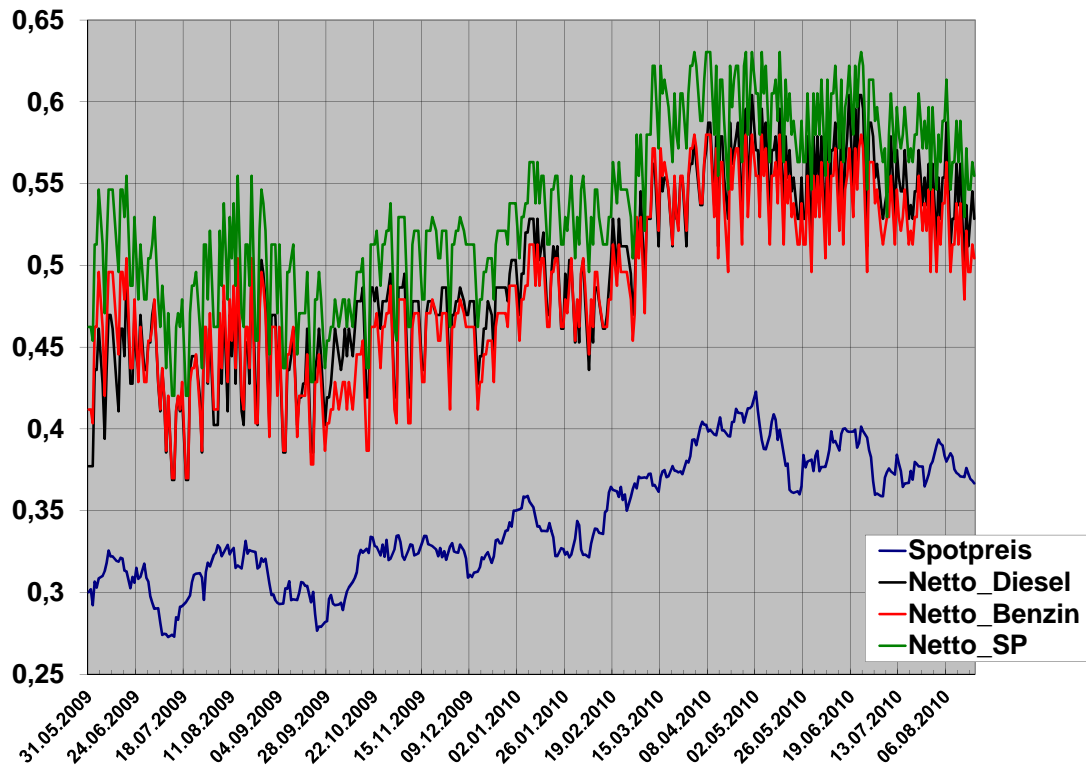


Abbildung 1: Bereinigte Benzin-, Diesel-, SuperPluspreise und Spotpreise in Euro

### 3.3 Einheitswurzel-Tests auf Stationarität

#### 3.3.1 Erweiterter Dickey-Fuller-Test

Mit Eviews wird der erweiterten Dickey-Fuller-Test (Augmented Dickey-Fuller-Test), ein Einheitswurzel-Test auf Stationarität durchgeführt. Die Ordnung  $p$  des AR(p)- Prozesses, den man dem Modell des erweiterten DF-Tests zugrunde legt, ist unbekannt. Mit Eviews wird  $p$  automatisch berechnet:

- für den Benzinpreis  $p = 12$ ,  $p_{max} = 16$
- für den Spotpreis  $p = 0$ ,  $p_{max} = 16$

Die Nullhypothese sagt aus, dass die Zeitreihe differenz-stationär ist. Die Alternativhypothese, dass eine stationäre Zeitreihe vorliegt. Es wird in Eviews der Test mit dem Trend und der Konstante gewählt, da in diesem Fall der deterministische Trend auch mitberücksichtigt wird, was ein genaueres Ergebnis bedeutet.

Der Augmented Dickey-Fuller Unit Root Test für den Benzinpreis in der Abbildung 2 zeigt, dass die  $t$ -Statistic von  $-2,853337$  größer als der kritische Wert bei 5 Prozent  $-3,424579$  ist. Das bedeutet, dass die Entscheidung im Bereich der Annahme von  $H_0$

liegt. Es gibt mindestens eine Einheitswurzel und somit liegt eine differenz-stationäre Zeitreihe vor.

Null Hypothesis: BENZIN has a unit root Exogenous: Constant, Linear Trend Lag Length: 12 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.853337	0.1795
Test critical values: 1% level	-3.988333	
5% level	-3.424579	
10% level	-3.135349	
*MacKinnon (1996) one-sided p-values.		

Abbildung 2: Erweiterter Dickey-Fuller Test für den Benzinpreis

Null Hypothesis: SPOT has a unit root Exogenous: Constant, Linear Trend Lag Length: 0 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-2.754462	0.2155
Test critical values: 1% level	-3.983684	
5% level	-3.422321	
10% level	-3.134016	
*MacKinnon (1996) one-sided p-values.		

Abbildung 3: Erweiterter Dickey-Fuller Test für den Spotpreis

Der Augmented Dickey-Fuller Unit Root Test für den Spotpreis in der Abbildung 3 zeigt, dass die **t-Statistic** von  $-2,754462$  größer als der kritische Wert bei 5 Prozent  $-3,422321$  ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_0$ . Die Zeitreihe hat wiederum mindestens eine Einheitswurzel und es liegt eine differenz-stationäre Zeitreihe vor.

Das Eliminieren der Nicht-Stationarität der Zeitreihen ist die Voraussetzung für eine gute Modellschätzung. Um aus einer nicht-stationären Zeitreihe eine stationäre Zeitreihe zu bekommen, werden die Zeitreihen mittels Differenzierung transformiert.

Null Hypothesis: D(BENZIN) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 11 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.275240	0.0000
Test critical values: 1% level	-3.988333	
5% level	-3.424579	
10% level	-3.135349	

\*MacKinnon (1996) one-sided p-values.

Abbildung 4: Erweiterter Dickey-Fuller Test für die erste Differenz der Benzinpreise

Null Hypothesis: D(SPOT) has a unit root		
Exogenous: Constant, Linear Trend		
Lag Length: 0 (Automatic - based on SIC, maxlag=16)		
	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-18.96369	0.0000
Test critical values: 1% level	-3.984496	
5% level	-3.422716	
10% level	-3.134249	

\*MacKinnon (1996) one-sided p-values.

Abbildung 5: Erweiterter Dickey-Fuller Test für die erste Differenz der Spotpreise

Der Augmented Dickey-Fuller Unit Root Test für den Benzinpreis in der Abbildung 4 zeigt, dass **t-Statistic** von  $-6,275240$  kleiner als der kritische Wert bei 5 Prozent  $-3,424579$  ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_1$ . Es liegt eine stationäre Zeitreihe vor.

Der P-Wert gleich  $0,000$  deutet auf die Signifikanz der Ergebnisse. Es ist die Wahrscheinlichkeit, einen so großen oder größeren (Absolut-)Wert der Teststatistik wie den beobachteten Wert bei Zutreffen der Nullhypothese zu realisieren. Die Nullhypothese besagt, dass die Variable keinen erklärenden Beitrag leistet, dass also der  $\beta$ -Koeffizient gleich Null ist. Die Alternativhypothese besagt, dass  $\beta$  ungleich Null ist. Die Nullhypo-

these wird verworfen, wenn p-Wert sehr klein ist (Hackl, 2005). Der Augmented Dickey-Fuller Unit Root Test für den Spotpreis in der Abbildung 5 zeigt, dass **t-Statistic** von  $-18,96369$  kleiner als der kritische Wert bei 5 Prozent  $-3,422716$  ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_1$ . Es liegt eine stationäre Zeitreihe vor. Der P-Wert gleich  $0,000$  deutet auf die Signifikanz der Ergebnisse.

### 3.3.2 Kwiatkowski-Phillips-Schmidt-Shin-Test

Der Kwiatkowski-Phillips-Schmidt-Shin Test wird angewendet, um das obige Ergebnis des erweiterten Dickey-Fuller Tests zu überprüfen. In diesem Einheitswurzel-Test sagt die Nullhypothese aus, dass die Zeitreihe stationär ist. Die Alternativhypothese besagt, dass es eine nicht-stationäre Zeitreihe vorliegt. In Eviews wird der Test mit dem Trend und mit der Konstante gewählt, da in diesem Fall der deterministische Trend auch mitberücksichtigt wird und somit ein genaueres Ergebnis liefert.

Null Hypothesis: BENZIN is stationary		
Exogenous: Constant, Linear Trend		
Bandwidth: 14 (Newey-West automatic) using Bartlett kernel		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.219196
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000

\*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Abbildung 6: KPSS-Test für den Benzinpreis

Das Ergebnis vom Kwiatkowski-Phillips-Schmidt-Shin Test für den Benzinpreis, dargestellt in der Abbildung 6, zeigt, dass **kpss-test-Statistic** von  $0,219196$  größer als der kritische Wert bei 5 Prozent  $0,146000$  ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_1$ . Es liegt eine nicht-stationäre Zeitreihe vor.

Der Kwiatkowski-Phillips-Schmidt-Shin Test für den Spotpreis in der Abbildung 7 zeigt, dass **kpss-test-Statistic** von  $0,166831$  größer als der kritische Wert bei 5 Prozent  $0,146000$  ist. Das bedeutet, die Entscheidung liegt im Bereich der Annahme von  $H_1$ . Es liegt eine nicht-stationäre Zeitreihe vor.

Null Hypothesis: SPOT is stationary	
Exogenous: Constant, Linear Trend	
Bandwidth: 15 (Newey-West automatic) using Bartlett kernel	
	LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic	0.166831
Asymptotic critical values*: 1% level	0.216000
5% level	0.146000
10% level	0.119000

\*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)

Abbildung 7: KPSS-Test für den Spotpreis

Das Ergebnis für die erste Differenz der Zeitreihe Benzinpreis, wie in der Abbildung 8 vorgestellt, zeigt, dass **kpss test statistics** von 0,050980 kleiner als **testcritical values** bei 5 Prozent 0,146000 ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_0$ . Die ersten Differenzen liefern die Stationarität.

Das Ergebnis für die erste Differenz der Zeitreihe Spotpreis in der Abbildung 9 zeigt, dass **kpss test statistics** von 0,048129 kleiner als **testcritical values** bei 5 Prozent 0,146000 ist. Das heißt, die Entscheidung liegt im Bereich der Annahme von  $H_0$ . Die ersten Differenzen liefern wiederum die Stationarität.



Null Hypothesis: D(BENZIN) is stationary		
Exogenous: Constant, Linear Trend		
Bandwidth: 31 (Newey-West automatic) using Bartlett kernel		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.050980
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Abbildung 8: KPSS-Test für die erste Differenz der Benzinpreise

Null Hypothesis: D(SPOT) is stationary		
Exogenous: Constant, Linear Trend		
Bandwidth: 8 (Newey-West automatic) using Bartlett kernel		
		LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic		0.048129
Asymptotic critical values*:	1% level	0.216000
	5% level	0.146000
	10% level	0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)		

Abbildung 9: KPSS-Test für die erste Differenz der Spotpreise

## 4 Periodizität und Saisonalität

### 4.1 Analyse der Zeitreihen auf Periodizität

In diesem Kapitel werden die Daten auf Periodizität und Saisonalität untersucht.

Um die Abhängigkeit der Nachbarelemente zu eliminieren, werden die ersten Differenzen der Zeitreihen gebildet. Danach werden die Korrelogramme für die ersten Differenzen mit der Formel für die Autokorrelationsfunktion berechnet (siehe 4.1). Für das Beschreiben des Verhaltens eines stochastischen Prozesses ist die AC-Funktion gut geeignet, da sie unabhängig von der Maßeinheit ist. Die graphische Darstellung der AC-Funktion nennt man *Korrelogramm* (Hackl, 2005). Die Größe der Spitze  $r_k$  im Korrelogramm charakterisiert die Korrelation zwischen dem Wert der Zeitreihe im Moment  $t$  und dem Wert der Zeitreihe im Moment  $t - k$ . Die Strichlinien bedeuten zwei Grenzen des doppelten Standardfehlers  $\pm 2/\sqrt{T}$  (Shumway & Stoffer, 2000).

$$\rho_k = \frac{\sum_{t=k+1}^T (y_t - \bar{y})(y_{t-k} - \bar{y}_{t-k}) / (T-k)}{\sum_{t=1}^T (y_t - \bar{y})^2 / T}, \quad \bar{y}_{t-k} = \sum y_{t-k} / T - k \quad (4.1)$$

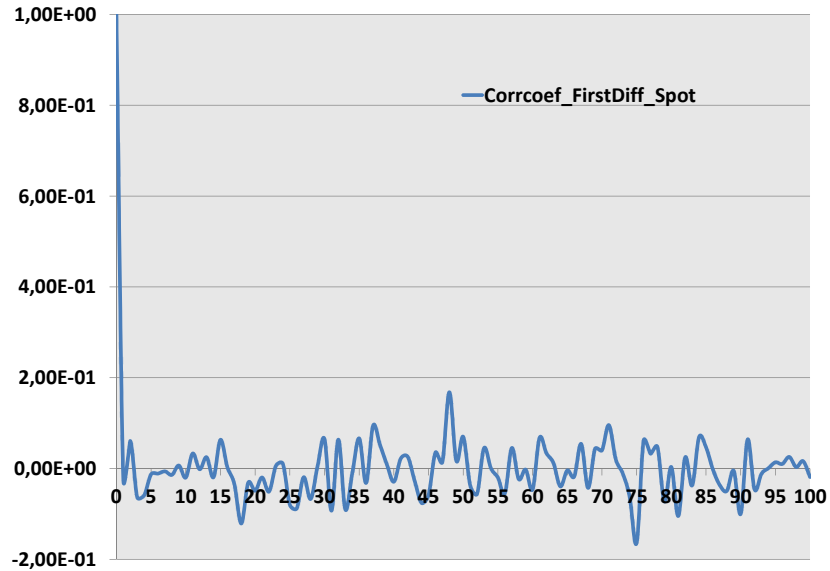


Abbildung 10: Korrelationsfunktion der ersten Differenzen der Spotpreise. Excel

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob
		1 -0.020	-0.020	0.1492	0.699
		2 0.057	0.057	1.3364	0.513
		3 -0.055	-0.053	2.4618	0.482
		4 -0.095	-0.101	5.7693	0.217
		5 0.005	0.007	5.7768	0.329
		6 -0.015	-0.007	5.8636	0.439
		7 -0.001	-0.013	5.8637	0.556
		8 -0.003	-0.011	5.8665	0.662
		9 -0.005	-0.004	5.8749	0.752
		10 0.018	0.016	5.9981	0.815
		11 0.013	0.013	6.0641	0.869
		12 -0.001	-0.004	6.0644	0.913
		13 0.012	0.011	6.1166	0.942
		14 -0.009	-0.003	6.1459	0.963
		15 0.078	0.080	8.4788	0.903
		16 0.017	0.022	8.5886	0.929
		17 -0.053	-0.062	9.6711	0.917
		18 -0.107	-0.107	14.056	0.725
		19 -0.043	-0.023	14.755	0.738
		20 -0.047	-0.040	15.620	0.740
		21 -0.010	-0.031	15.658	0.789
		22 -0.017	-0.038	15.775	0.827
		23 -0.009	-0.020	15.809	0.863
		24 -0.016	-0.027	15.913	0.891
		25 -0.060	-0.073	17.300	0.871
		26 -0.063	-0.082	18.857	0.842
		27 -0.025	-0.031	19.110	0.866
		28 -0.049	-0.056	20.055	0.862
		29 0.021	0.003	20.226	0.886
		30 0.045	0.032	21.014	0.887
		31 -0.104	-0.125	25.342	0.752
		32 0.053	0.038	26.465	0.743
		33 -0.047	-0.008	27.342	0.745
		34 -0.002	-0.014	27.343	0.784
		35 0.045	0.028	28.152	0.788
		36 -0.011	-0.010	28.205	0.820
		37 0.096	0.085	31.917	0.706
		38 0.038	0.048	32.515	0.721
		39 -0.005	-0.015	32.526	0.758
		40 -0.016	-0.016	32.627	0.790
		41 -0.018	0.012	32.765	0.817
		42 0.025	0.029	33.025	0.838
		43 -0.025	-0.043	33.289	0.857
		44 -0.072	-0.113	35.435	0.818
		45 -0.069	-0.117	37.429	0.781
		46 0.029	0.029	37.781	0.800
		47 -0.003	-0.027	37.785	0.829
		48 0.155	0.106	47.785	0.482
		49 0.033	0.003	48.230	0.504
		50 0.062	0.040	49.844	0.480
		51 -0.030	-0.038	50.232	0.504
		52 -0.050	-0.060	51.302	0.501
		53 0.028	0.013	51.636	0.527
		54 -0.019	0.006	51.782	0.560
		55 -0.027	-0.003	52.093	0.586
		56 -0.027	-0.019	52.397	0.612
		57 0.028	0.035	52.743	0.635
		58 -0.021	-0.026	52.928	0.664
		59 -0.025	-0.037	53.194	0.688
		60 -0.012	0.016	53.258	0.719
		61 0.037	0.041	53.843	0.730
		62 0.054	0.032	55.118	0.720
		63 0.026	0.010	55.421	0.740
		64 -0.026	-0.049	55.714	0.760
		65 -0.015	-0.014	55.812	0.785
		66 0.012	0.060	55.871	0.808
		67 0.003	0.015	55.876	0.832
		68 -0.021	-0.002	56.073	0.849
		69 0.029	0.013	56.445	0.861
		70 0.022	0.018	56.662	0.875
		71 0.072	0.074	59.000	0.845
		72 0.014	-0.010	59.092	0.862
		73 -0.010	0.002	59.141	0.880
		74 -0.041	-0.005	59.898	0.882
		75 -0.108	-0.097	65.250	0.782
		76 0.064	0.056	67.130	0.756
		77 0.021	0.019	67.335	0.776
		78 0.031	-0.018	67.785	0.789
		79 -0.063	-0.046	69.649	0.765
		80 -0.047	-0.044	70.693	0.762

Abbildung 11: Korrelogramm der ersten Differenzen der Spotpreise. Eviews

Die Korrelationsfunktion der ersten Differenzen der Spotpreise in Excel in der Abbildung 10 sowie das Korrelogramm in Eviews 11 (Abbildung 11) zeigen zwei Spitzen an den Verzögerungsstellen 48 und 75. Die Werte dieser Spitzen befinden sich aber in den Grenzen des Standardfehlers (in den Grenzen der Strichlinie). Deswegen kann man diese Spitzen außer Acht lassen und behaupten, dass es keine klare Periodizität der Spotpreise vorhanden ist.

Die Korrelogramme der ersten Differenzen der Diesel-, Benzin-, SuperPluspreise sind in den Abbildungen 13, 14 und 15 dargestellt. Es ist interessant zu beobachten, dass alle drei Korrelogramme große Spitzen an den durch 6 teilbaren Verzögerungsstellen aufweisen. Zu bemerken ist, dass diese Spitzen bedeutend größer als der doppelte Standardfehler sind. Diese Tatsache besagt, dass es eine Periodizität mit einer Verzögerung 6 vorhanden ist. Da die Daten nur 6 Tage die Woche beinhalten, das heißt alle Wochentage außer Samstag, muss man schließen, dass es sich hier um eine wöchentliche Periodizität handelt. Die in Excel generierte Korrelationsfunktion am Beispiel der Dieselpreise verdeutlicht das gleiche Ergebnis einer Verzögerung 6 (Abbildung 12).

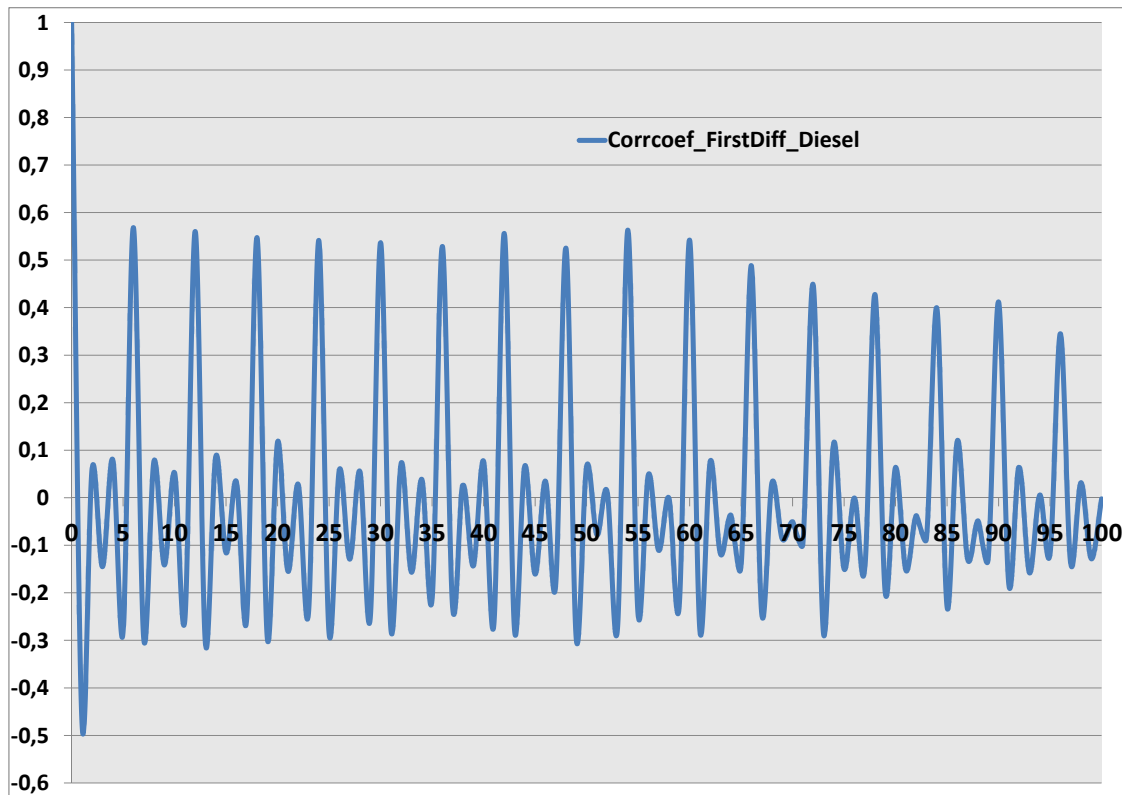


Abbildung 12: Korrelationsfunktion der ersten Differenzen der Dieselpreise. Excel

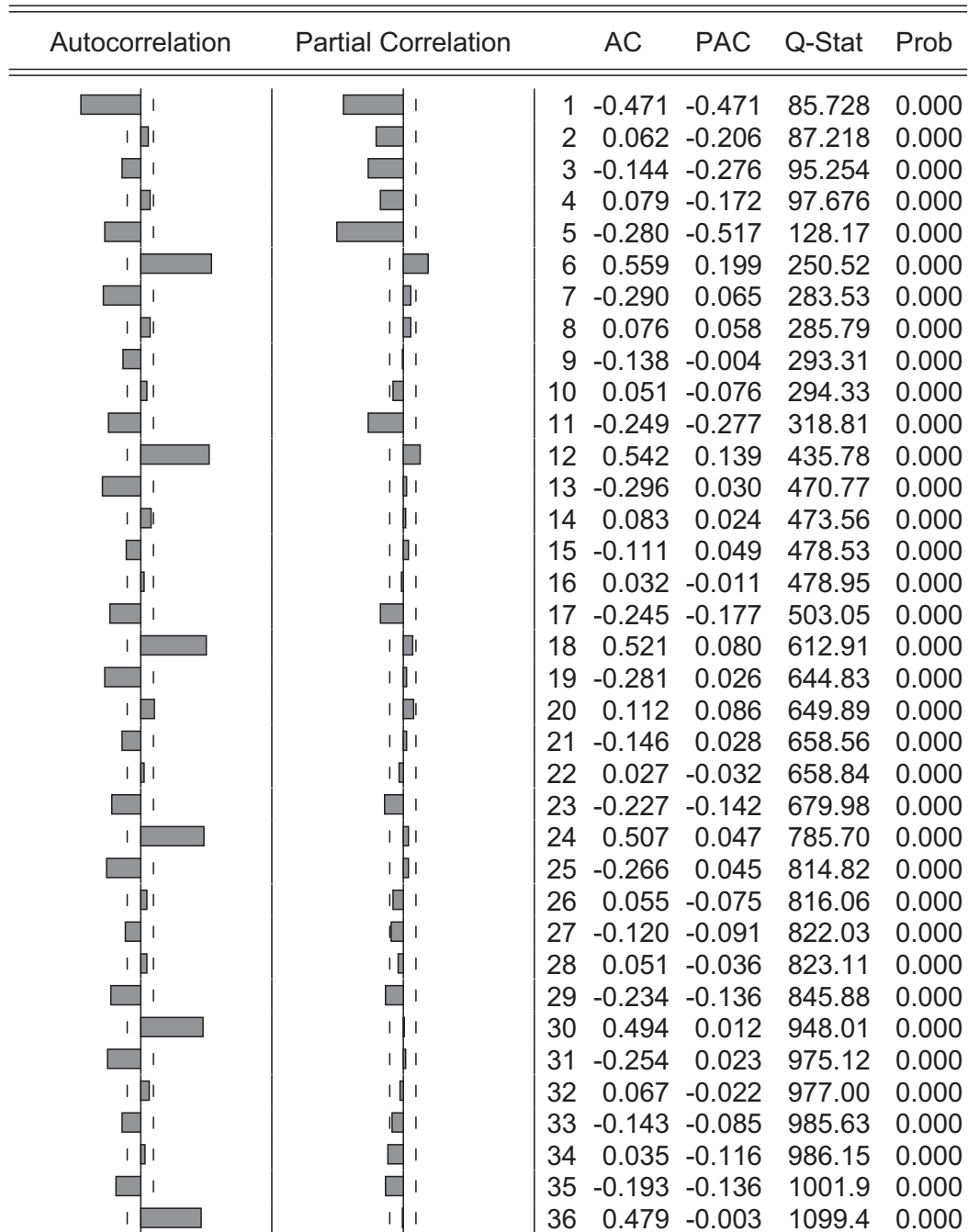


Abbildung 13: Korrelogramm der ersten Differenzen der Dieselpreise. Eviews

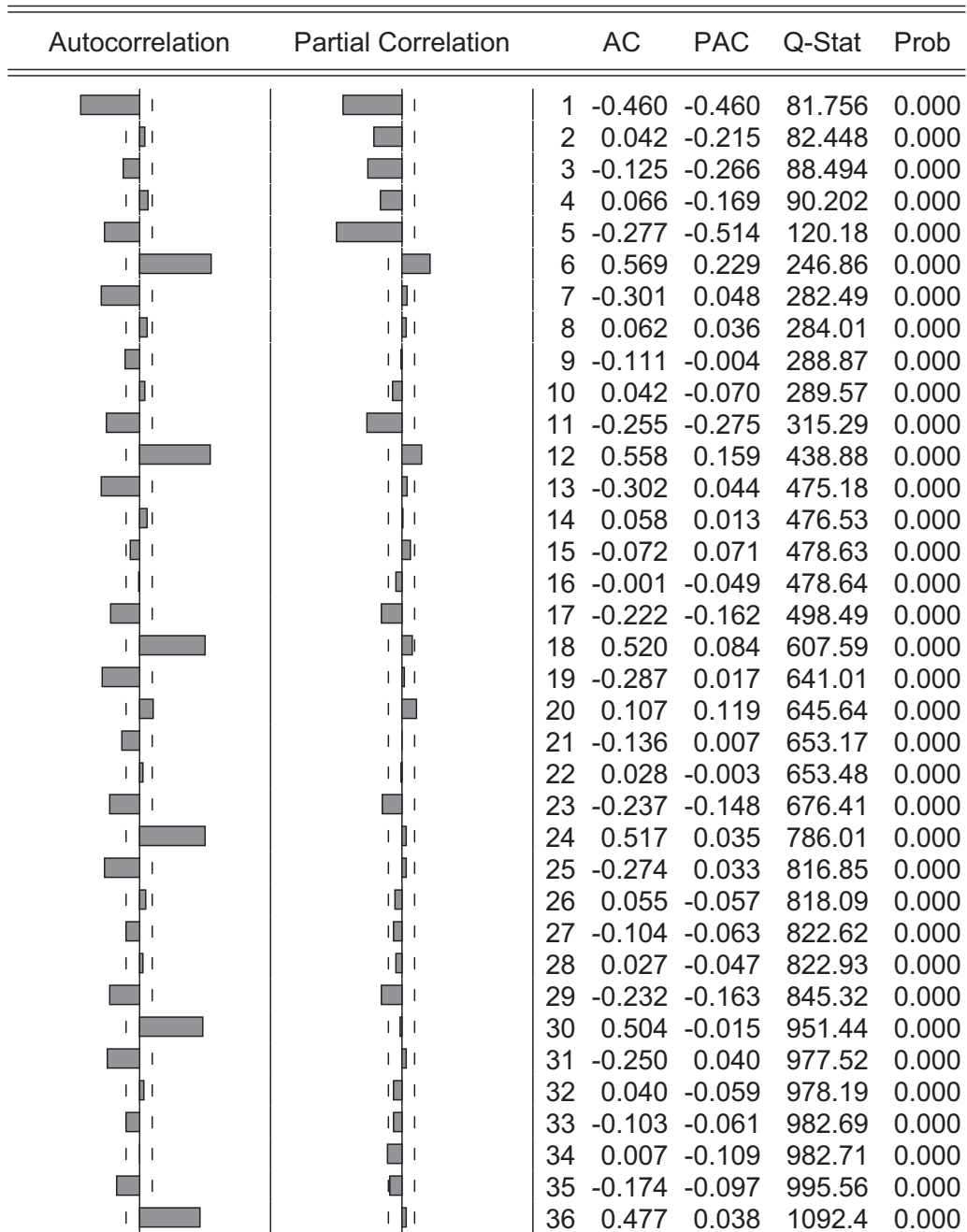


Abbildung 14: Korrelogramm der ersten Differenzen der Benzinpreise. Eviews

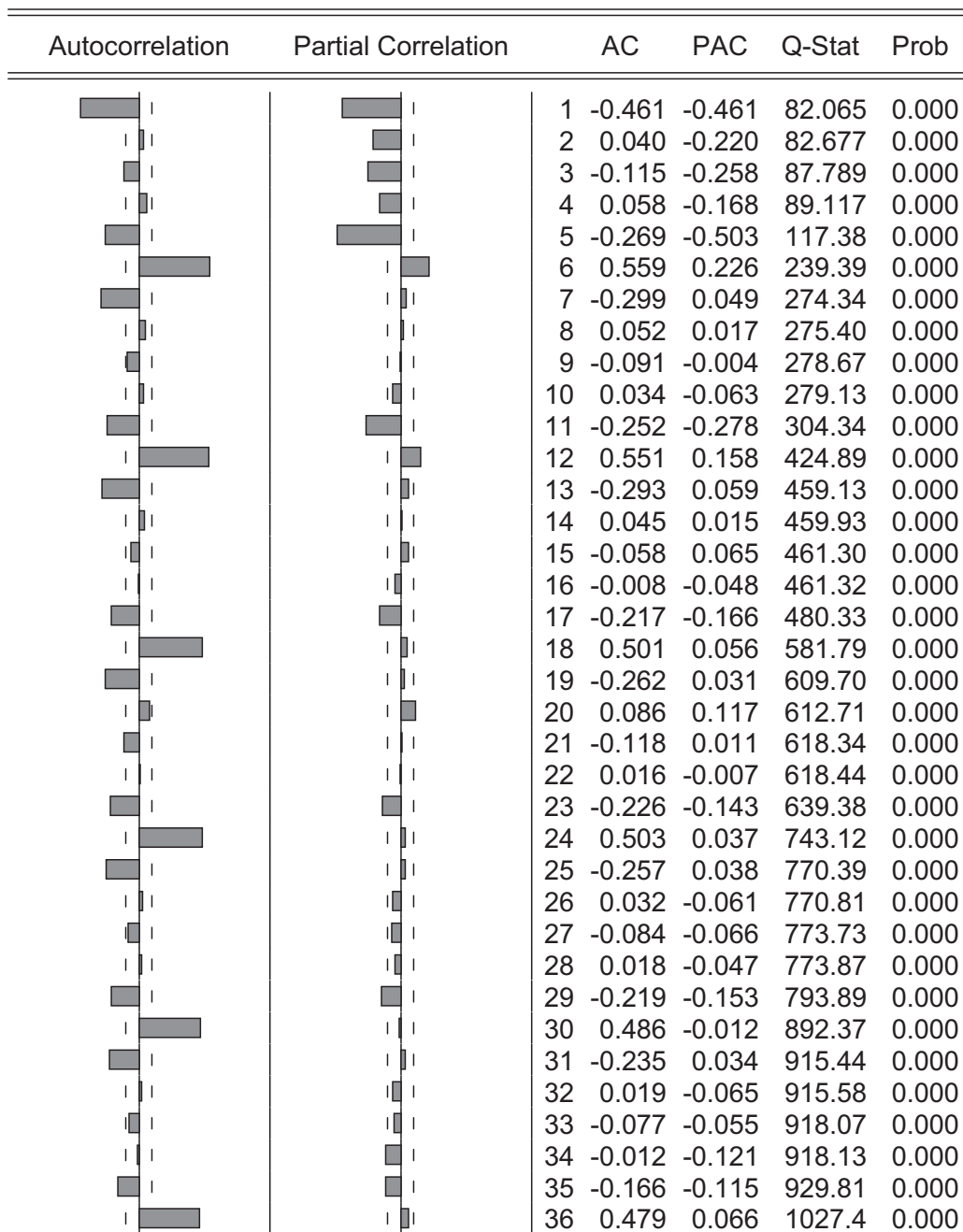


Abbildung 15: Korrelogramm der ersten Differenzen der SuperPluspreise. Eviews

## 4.2 Saisonale Zerlegung der Zeitreihen

Bei der Analyse der Zeitreihe auf Periodizität ist es von Vorteil, die Zeitreihe in einige Komponenten zu teilen. Man muss aus den Originaldaten Komponenten eliminieren, die dazu führen, dass es große Spitzen in den Korrelogrammen gibt. Sie spielen in der Periodizität eine dominante Rolle, da diese Spitzen auch andere Spitzen unterdrücken können, die durch die anderen saisonalen und periodischen Komponenten hervorgerufen werden.

Bei der Modellierung besteht die Wahl zwischen einem additiven oder einem multiplikativen Modell. Die Idee dieser Modelle besteht darin, dass die Prognosen nicht nur auf den Beobachtungen selbst beruhen, sondern auch auf den Zeitverzögerungen (Lags), so dass man die saisonale und die Trendkomponenten unabhängig voneinander schätzen kann.

Die Idee von einem additiven und multiplikativen Modell kann man am Beispiel des Spielzeugverkaufs darstellen. Der Verkauf an Spielzeugen steigt im Dezember vor Weihnachten oder während der Sommerferien. Das heißt, es wird jedes Jahr eine große Spitze im Dezember geben und eine kleinere im Zeitraum der Sommerferien. Diese saisonale Abhängigkeit wiederholt sich jedes Jahr. An diesem Beispiel kann man den Unterschied zwischen der additiven und multiplikativen Komponenten beobachten. Zum Beispiel steigt im Dezember der Verkauf an Spielzeugen um 3 Millionen Euro. Um die saisonale Komponente zu berücksichtigen, kann man in die Prognose für den Dezember 3 Millionen Euro hinzufügen. In diesem Fall wäre die Saisonalität additiv.

Der Verkauf an Spielzeugen kann aber auch im Dezember um 40 Prozent steigen, das bedeutet mit 1,4 multipliziert. Wenn der Durchschnittsumsatz (in Euro) an Spielzeugen gering ist, so ist der absolute Zuwachs der Verkäufe im Dezember auch gering (der prozentuale Zuwachs ist konstant). Wenn Spielzeuge gut verkauft werden, dann fällt der prozentuale Zuwachs (in Euro) deutlich größer aus. Der Umsatz steigt 1,4 mal und es liegt eine multiplikative Saisonalität vor. In diesem Fall wäre der Parameter der multiplikativen saisonalen Komponente gleich 1,4. Graphisch gesehen, wäre der Unterschied zwischen den beiden Saisonalitäten folgenderweise zu beurteilen: im additiven Modell werden die saisonalen Schwankungen konstant sein, deren Größe vom allgemeinen Niveau der Werte der Reihe nicht abhängig ist, in einem multiplikativen Modell wird aber die Größe der Schwankungen vom allgemeinen Niveau der Werte der Reihe abhängen.

### Additives Modell

- $X_t = T_t + Z_t + S_t + I_t$

### Multiplikatives Modell

- $X_t = T Z_t + S_t + I_t$



Die Zeitreihe besteht aus vier verschiedenen Komponenten: der saisonalen Komponente  $S_t$  im Moment  $t$ , des Trends  $T_t$ , der zyklischen Komponente  $C_t$  und der zufälligen, nicht regulären Komponente oder Fluktuationen  $I_t$ . Die saisonale Komponente hat einen regulären (saisonalen) Verlauf, die zyklische hat aber eher einen längerfristigen Effekt, der sich zyklisch ändern kann.  $X_t$  bedeutet den Wert der Zeitreihe im Moment  $t$ .

Um die periodische (additive und multiplikative) Komponente zu bestimmen, muss man zuerst die Originalzeitreihe glätten, um irreguläre Schwankungen zu eliminieren. Mit diesem Ziel wird die Methode der gleitenden Durchschnitte angewendet. Dies geschieht mit der Formel 4.2. Die erzeugte Reihe wird als MA6 bezeichnet.

$$MA_t(k) = \frac{x_{t-(k-1)} + x_{t-(k-2)} + \dots + x_t}{k} \Rightarrow MA_t(6) = \frac{x_{t-5} + x_{t-4} + \dots + x_t}{6} \quad (4.2)$$

#### 4.2.1 Multiplikatives und additives Modell

Im weiteren werden die Kraftstoffpreise am Beispiel des Dieselpreises untersucht. Im Fall von einem multiplikativen Modell nimmt man das Verhältnis der Originalzeitreihe der Dieselpreise zum gleitenden Durchschnitt der Zeitreihe der Dieselpreise  $TS/MA6$ . Die Variationen der Durchschnittswerte von  $TS/MA6$  für jeweils jeden Wochentag (außer Samstag) für jeweils jeden Monat des Jahres sind graphisch in der Abbildung 16 dargestellt. Es ist interessant zu beobachten, wie sich die durchschnittlichen monatlichen Werte von  $TS/MA6$  im Zeitverlauf verändern.

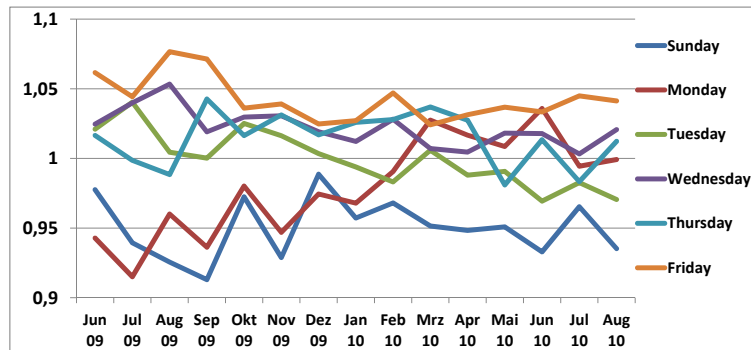


Abbildung 16: TS/MA6. Monatlich. Excel

Beim Vergleich der Jahre 2009 und 2010 fällt auf, dass die mittleren Wochentagspreisschwankungen im Jahr 2009 deutlich größer als deren Schwankungen im Jahr 2010. Es ist auch interessant, dass die Montagspreisänderungen parallel zu den Sonntagspreisänderungen im Jahr 2009 verlaufen, indem im Jahr 2010 die Montagsschwankungen

eher einen negativen Trend haben und die Sonntagspreisänderungen eher einen positiven Trend. Die Dienstagspreisänderungen weisen über den ganzen Zeitraum eher einen negativen Trend auf.

Die Werte von TS/MA6 liegen generell zwischen Sonntags- und Freitagspreisen. So ist an der Graphik deutlich zu sehen, dass die Freitagspreise viel höher als die Durchschnittspreise sind und das ca. um 2,5 bis 7,5 Prozent über dem Durchschnitt. Die prozentualen Sonntagspreisschwankungen liegen ca. 1 bis 8 Prozent unter dem Durchschnitt.

Auf diese Weise entsteht ein so genannter Kanal, in dem sich die Werte von TS/MA6 verändern. Man merkt, dass die Breite dieses Kanals im Winter enger als im Sommer ist. Das heißt, dass die Variabilität in den Preisschwankungen im Sommer größer ist als im Winter. Deswegen ist es sinnvoll, die Durchschnittswerte von TS/MA6 für jeweils jeden Wochentag (außer Samstag) für jeweils jeden Zeitabschnitt des Jahres darzustellen, so die Abbildung 17.

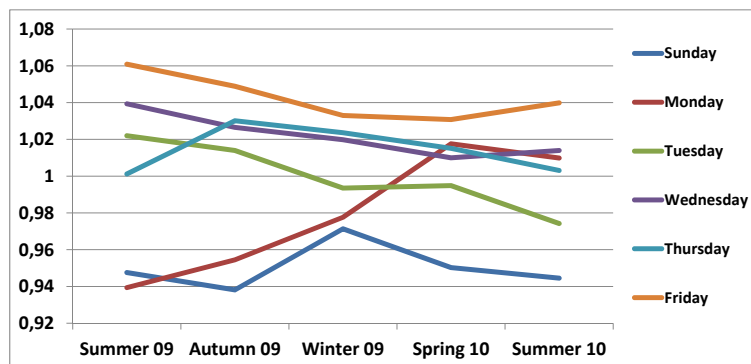


Abbildung 17: TS/MA6. Saisonweise. Excel

Im Fall mit einem additiven Modell werden die Differenzen der Originalzeitreihe der Dieselpreise zum gleitenden Durchschnitt der Zeitreihe der Dieselpreise TS-MA6 untersucht. Die Variationen der Durchschnittswerte von TS-MA6 für jeweils jeden Wochentag (außer Samstag) für jeweils jeden Monat des Jahres sind graphisch in der Abbildung 18 dargestellt. Es ist interessant, wie sich die durchschnittlichen monatlichen Werte von TS-MA6 im Zeitverlauf verändern.

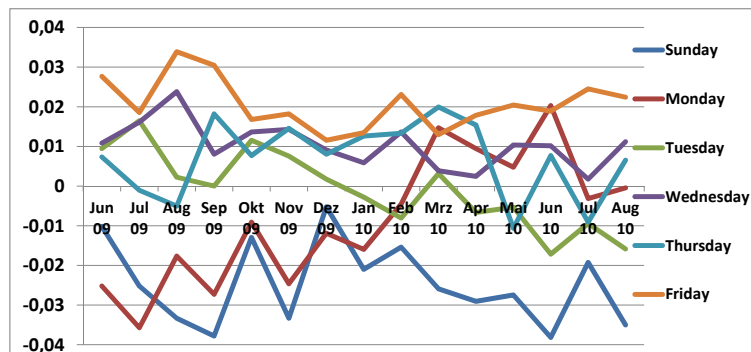


Abbildung 18: TS-MA6. Monatlich. Excel

Man sieht, dass das Bild auf dieser Graphik den oben beschriebenen Verlauf der Preisschwankungen des multiplikativen Modells wiederholt. Die Durchschnittswerte von TS-MA6 für jeweils jeden Wochentag (außer Samstag) für jeweils jeden Zeitabschnitt des Jahres in der Abbildung 19 weisen wieder das gleiche Verhalten wie die der Durchschnittswerte im multiplikativen Modell auf. Deswegen kann keine Aussage bezüglich der Art des Modells getroffen werden. Die beiden Modelle haben aber gezeigt, dass der Freitagspreis im Wochenverlauf am höchsten ist. Dies bestätigt die erste aufgestellte Hypothese.

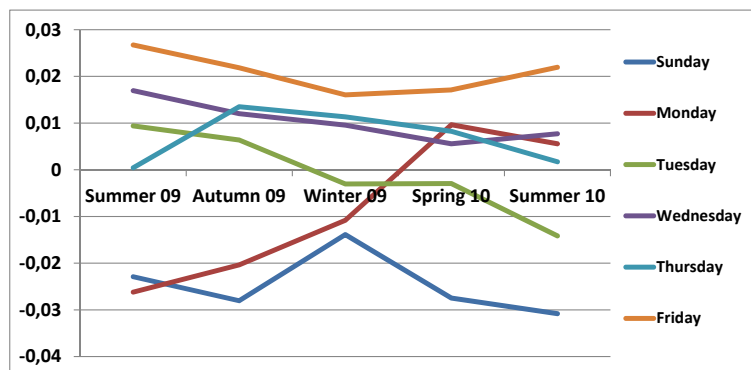


Abbildung 19: TS-MA6. Saisonal. Excel

Anhand der beiden Graphiken in den Abbildungen 20 und 21 wird zwischen den beiden Modellen entschieden. Hier wird die Bildung der Differenz zwischen den durch-

schnittlichen monatlichen Freitags- und Sonntagspreisen vollzogen. Man sieht an der graphischen Abbildung für additives Modell eine klarere saisonale Abhängigkeit, weil der Durchschnittswert der Breite des Kanals im Sommer 2009 dem Durchschnittswert der Breite des Kanals im Sommer 2010 im additiven Modell ähnlicher ist. Deswegen wird das additive Modell dem multiplikativen Modell vorgezogen.

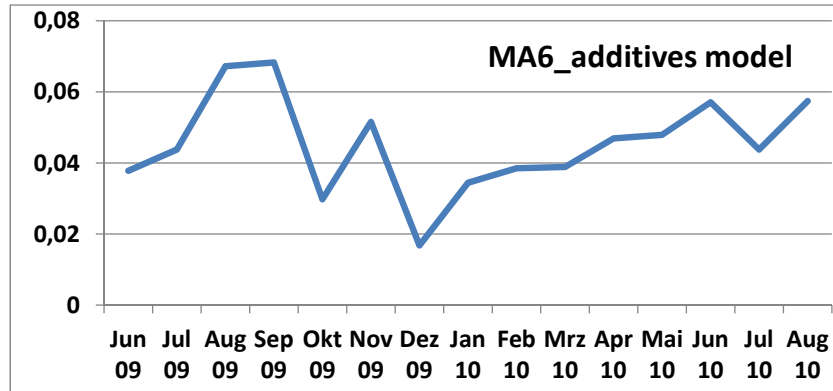


Abbildung 20: Additives Modell. Excel

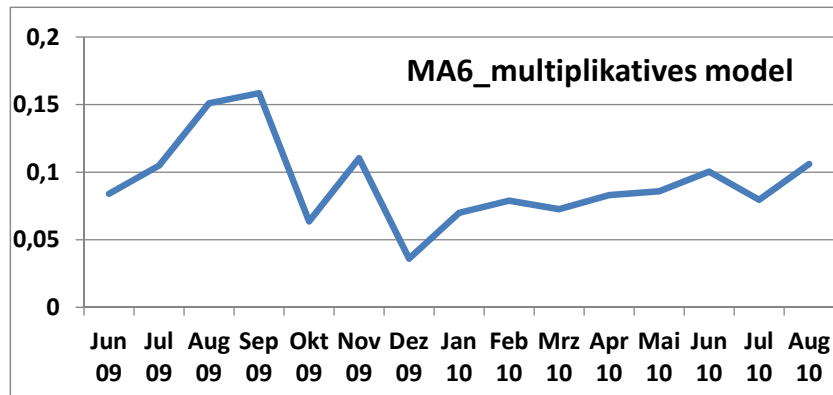


Abbildung 21: Multiplikatives Excel

### 4.2.2 Anwendung des additiven Modells

Die additive Komponente muss jetzt gebildet werden. Mithilfe der Tabelle in der Abbildung 22 bestimmt man den theoretischen Prozess `TeorDiffDiesel` der saisonalen Komponente. In jeder Zelle befindet sich der durchschnittliche Wert der Differenz `TS-MA6` für jeden Wochentag für jede Jahreszeit.

	<i>Sommer</i>	<i>Herbst</i>	<i>Winter</i>	<i>Frühling</i>
Sonntag	-0,02678	-0,02844	-0,01443	-0,02747
Montag	-0,01102	0,02068	-0,01120	0,00926
Dienstag	-0,00163	0,00646	-0,00301	-0,00312
Mittwoch	0,01283	0,01217	0,00926	0,00592
Donnerstag	0,00058	0,01357	0,01141	0,00678
Freitag	0,02480	0,02154	0,01583	0,01734

Abbildung 22: Durchschnittswerte der Differenz `TS-MA6`

Anhand der Tabelle wird die Matrix  $D_{ij}$  gebildet, wo die Werte  $i$  der Nummer der Zeile (1 bis 6) und die Werte  $j$  der Nummer der Spalte (1 bis 4) entsprechen. Die saisonalen Effekte werden in einem Modell mittels Dummy-Variablen beschrieben. Es werden die Dummy-Variablen  $A_{jT}$  und  $B_{iT}$  eingeführt:

$$A_{1T} \begin{cases} 1, & \text{wenn Monat}(T) = \text{Dezember, Januar oder Februar} \\ 0, & \text{wenn Monat}(T) \neq \text{Dezember, Januar oder Februar} \end{cases}$$

$$A_{2T} \begin{cases} 1, & \text{wenn Monat}(T) = \text{März, April, Mai} \\ 0, & \text{wenn Monat}(T) \neq \text{März, April, Mai} \end{cases}$$

$$A_{3T} \begin{cases} 1, & \text{wenn Monat}(T) = \text{Juni, Juli, August} \\ 0, & \text{wenn Monat}(T) \neq \text{Juni, Juli, August} \end{cases}$$

$$A_{4T} \begin{cases} 1, & \text{wenn Monat}(T) = \text{September, Oktober, November} \\ 0, & \text{wenn Monat}(T) \neq \text{September, Oktober, November} \end{cases}$$

$$B_{1T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Sonntag} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Sonntag} \end{cases}$$

$$B_{2T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Montag} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Montag} \end{cases}$$

$$B_{3T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Dienstag} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Dienstag} \end{cases}$$

$$B_{4T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Mittwoch} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Mittwoch} \end{cases}$$

$$B_{5T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Donnerstag} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Donnerstag} \end{cases}$$

$$B_{6T} \begin{cases} 1, & \text{wenn Wochentag}(T) = \text{Freitag} \\ 0, & \text{wenn Wochentag}(T) \neq \text{Freitag} \end{cases}$$

Auf diese Weise wird die saisonale Komponente  $S_t$  ermittelt, die mit der Formel 4.3 berechnet wird. Danach wird die saisonale Komponente  $S_t$  von der Originalzeitreihe der Dieselpreise subtrahiert. Die neue Zeitreihe **TS\_Teordiffdiesel** ist saisonbereinigt. Zu beachten ist, dass im Korrelogramm der ersten Differenzen der saisonbereinigten Zeitreihe **TS\_Teordiffdiesel** in der Abbildung 23 keine Spitzen an den durch 6 teilbaren Verzögerungsstellen zu sehen sind. Das Ziel der saisonalen Zerlegung wurde erreicht. Das ARMA-Modell wird im weiteren auf die neue saisonbereinigte Zeitreihe angewendet.

$$S_t = \sum_{j=1}^4 \sum_{i=1}^6 B_{it} A_{jt} D_{ij} \quad (4.3)$$

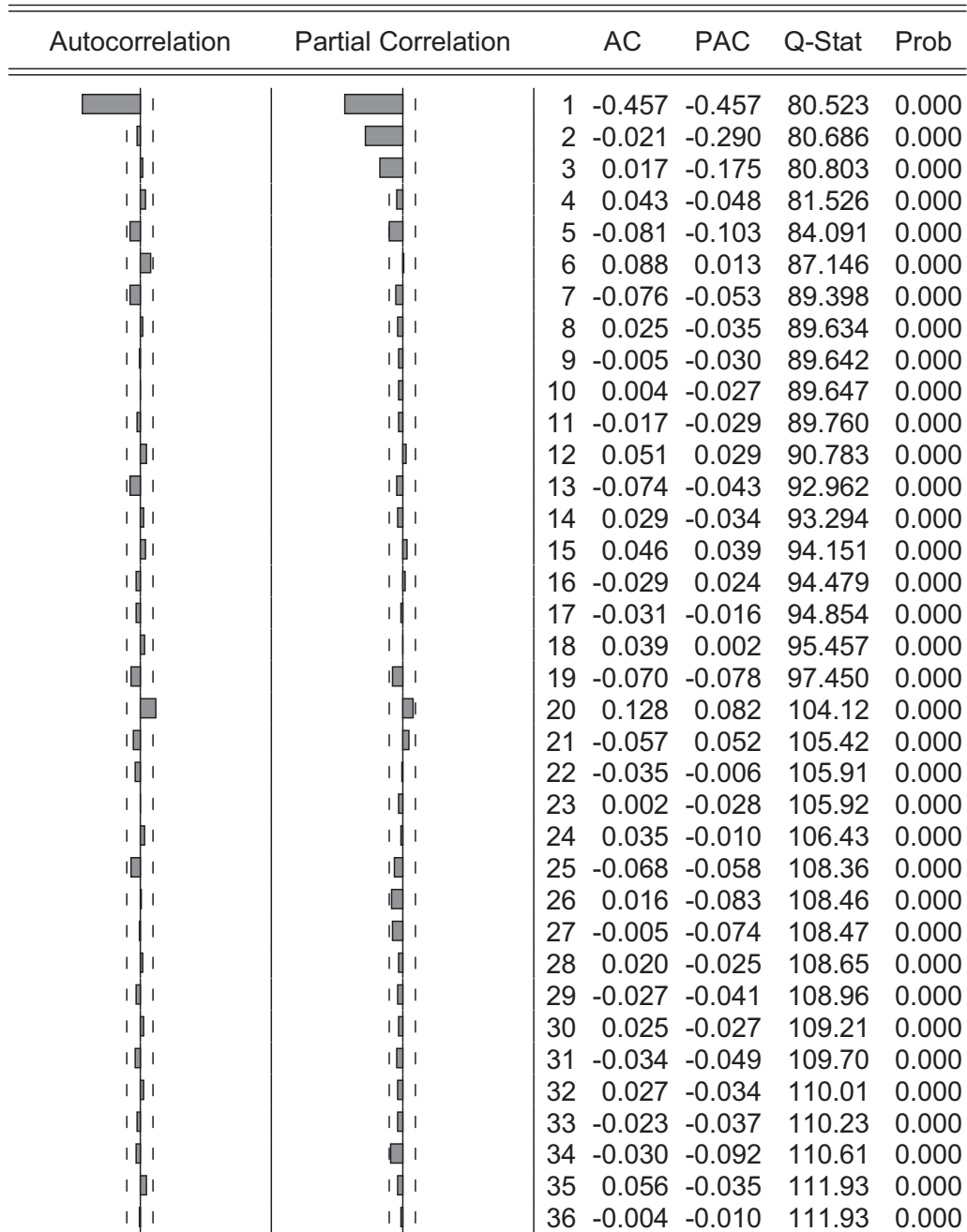


Abbildung 23: Korrelogramm der ersten Differenzen von TS\_Teordiffdiesel

## 5 Modellschätzung

### 5.1 Modellbildung der Spotpreise

Am Anfang wird das Modell für den Spotpreis gebildet. Für die Bildung der ARMA-Modelle werden die ersten und die zweiten Differenzen untersucht. Beide Zeitreihen weisen die Stationarität auf und sind für die Modellbildung geeignet.

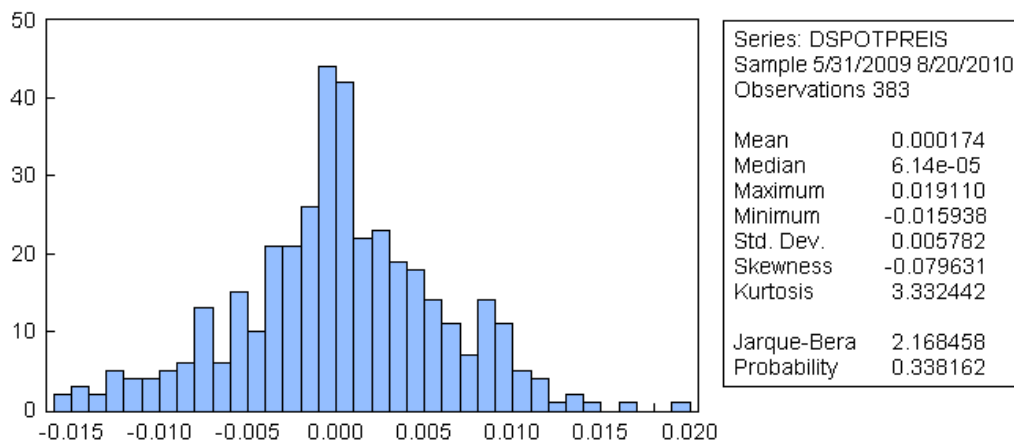


Abbildung 24: Histogramm der Verteilung der ersten Differenz der Spotpreise

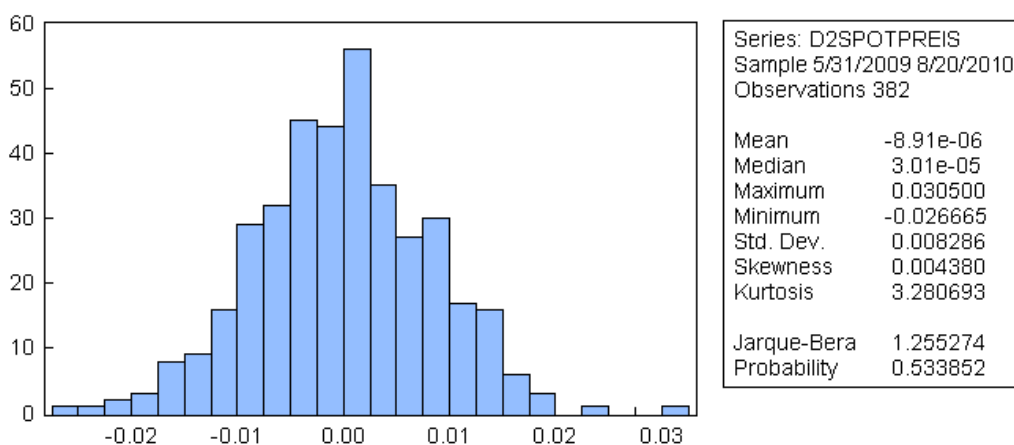


Abbildung 25: Histogramm der Verteilung der zweiten Differenz der Spotpreise

Die Histogramme in den Abbildungen 24 und 25 zeigen die empirischen Verteilungen der ersten und der zweiten Differenzen der Spotpreise. Man sieht, dass die Verteilungen



symmetrisch sind. Die beiden Verteilungen werden auf die Schiefe **Skewness** und die Wölbung **Kurtosis** der Dichtefunktion geprüft und mit der Schiefe und der Wölbung der Normalverteilung verglichen. Es wird diejenige Zeitreihe für die Modellbildung bevorzugt, deren Wölbung der Dichtefunktion einer Normalverteilung ähnlicher ist und deren Schiefe am geringsten ist (Weigend & Gershenfeld, 1994).

Die Wölbung  $\beta_2$  misst die Steilheit der Dichtedunktion. Die Normalverteilung hat die Wölbung von  $\beta_2 = 3$ . Die Schiefe  $v$  zeigt, ob die Verteilung rechts- oder links-schief ist. Die Normalverteilung weist eine Schiefe von Null auf. Die beiden Maßzahlen werden nach den Formeln  $\beta_2 = \tilde{\mu}_3 / \tilde{\sigma}^3$  und  $v = \tilde{\mu}_4 / \tilde{\sigma}^4$  mit  $\tilde{\sigma} = \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 / N}$ ,  $\tilde{\mu}_3 = (X_i - \bar{X})^3 / N$ ,  $\tilde{\mu}_4 = (X_i - \bar{X})^4 / N$  und  $\bar{X} = \sum_{i=1}^n X_i / N$  berechnet (Agung, 2010).

**Skewness** der zweiten Differenzen hat den Wert 0,004380 und ist näher an 0 als **Skewness** der ersten Differenzen mit dem Wert von -0,079631. **Kurtosis** der zweiten Differenzen hat den Wert 3,280693 und ist näher an 3 als **Kurtosis** der ersten Differenzen mit dem Wert von 3,332442. Aus diesem Grund wird die zweite Differenz der Spotpreise für den ARMA-Modellbau bevorzugt.

Dependent Variable: N_D2SPOTPREIS Method: Least Squares Date: 12/30/11 Time: 13:26 Sample (adjusted): 6/02/2009 8/20/2010 Included observations: 382 after adjustments Convergence achieved after 13 iterations MA Backcast: 6/01/2009				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	7.77E-06	3.00E-06	2.594081	0.0099
MA(1)	-0.995859	0.003449	-288.7182	0.0000
R-squared	0.509672	Mean dependent var	-3.92E-09	
Adjusted R-squared	0.508381	S.D. dependent var	0.008286	
S.E. of regression	0.005810	Akaike info criterion	-7.453421	
Sum squared resid	0.012825	Schwarz criterion	-7.432765	
Log likelihood	1425.603	Hannan-Quinn criter.	-7.445226	
F-statistic	394.9908	Durbin-Watson stat	2.037616	
Prob(F-statistic)	0.000000			
Inverted MA Roots	1.00			

Abbildung 26: Modellierung des MA-Prozesses für die zweite Differenz der Spotpreise

Die zweite Differenz der Spotpreise wird mit dem ARMA-Modell (2,0,1) approximiert, das heißt mit dem MA-Prozess ( $p = 0$ ). Aus der Abbildung 26 ist ersichtlich, dass der MA-Prozess mit der Ordnung 1 für die Beschreibung der zweiten Differenzen der Spotpreise signifikant ist, denn die p-Werte von MA(1) und C nah an 0 sind. **Durbin-Watson-Stat**

mit dem Wert von 2,037616 ist nah an 2.

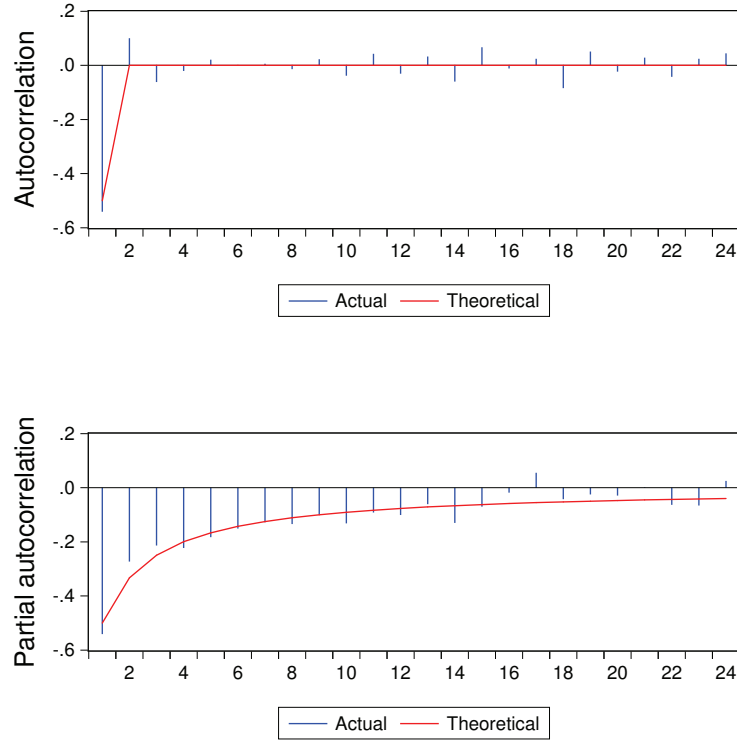


Abbildung 27: Autocorrelogramm der zweiten Differenzen der Spotpreise

In der Abbildung 27 sieht man eine Übereinstimmung zwischen dem Korrelogramm für den MA-Prozess der 1. Ordnung und dem Korrelogramm für die zweiten Differenzen der Spotpreise. Dies bedeutet, es liegt ein hoch signifikantes Modell vor (Weigend & Gershenfeld, 1994).

Die Zeitreihe der zweiten Differenzen der Spotpreise wird als  $Z_t$  bezeichnet. Das gebildete Modell für diese Zeitreihe sieht folgenderweise aus:

$$Z_t = \mu + c + \alpha \cdot \epsilon_{t-1} + \epsilon_t = -0,00000891 + 0,00000777 - 0,995859\epsilon_{t-1} + \epsilon_t \quad (5.1)$$

Mittels Integralrechnung wird das Modell für die Originalreihe der Spotpreise  $X_t$  gebildet:

$$X_t = X_0 + t \cdot (X_1 - X_0) + \sum_{\tau=1}^t \left( \sum_{\zeta=\tau}^{t-1} (\mu + \alpha \cdot \epsilon_{\zeta} + \epsilon_{\tau+\zeta}) \right) \quad (5.2)$$

Man setzt  $\alpha$  und  $\mu$  in die Formel 5.2 ein und bekommt folgendes Modell:

$$X_t = X_0 + t \cdot (X_1 - X_0) + \sum_{\tau=1}^t \left( \sum_{\zeta=\tau}^{t-1} (-0,00000891 + 0,995859 \cdot \epsilon_{\zeta} + \epsilon_{\tau+\zeta}) \right) \quad (5.3)$$

Dies ist das gesuchte Modell für die Originalzeitreihe der Dieselpreise, das man für die Prognose verwenden kann. Wenn man über die Beobachtungswerte der vorherigen Perioden  $t=0,1,2$  verfügt, kann man eine Vorhersage für die Werte von  $X_t$  ab der Periode  $t=5$  machen.

## 5.2 Modellbildung der Dieselpreise

Im Weiteren wird nach dem Modell für die Kraftstoffpreise am Beispiel der Dieselpreise gesucht. Für die Modellbildung nimmt man die zweite Differenz der saisonbereinigten Zeitreihe der Dieselpreise.

Dependent Variable: N_D2TEORDIESEL Method: Least Squares Date: 12/30/11 Time: 16:20 Sample (adjusted): 6/05/2009 8/20/2010 Included observations: 379 after adjustments Convergence achieved after 15 iterations MA Backcast: 6/04/2009				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.18E-05	5.34E-06	11.55575	0.0000
AR(1)	-0.652585	0.050772	-12.85317	0.0000
AR(2)	-0.393598	0.057060	-6.897959	0.0000
AR(3)	-0.172214	0.050773	-3.391823	0.0008
MA(1)	-0.992617	0.003104	-319.7419	0.0000
R-squared	0.760435	Mean dependent var	-7.63E-05	
Adjusted R-squared	0.757873	S.D. dependent var	0.038206	
S.E. of regression	0.018800	Akaike info criterion	-5.096853	
Sum squared resid	0.132181	Schwarz criterion	-5.044907	
Log likelihood	970.8537	Hannan-Quinn criter.	-5.076239	
F-statistic	296.7907	Durbin-Watson stat	2.016632	
Prob(F-statistic)	0.000000			
Inverted AR Roots	-.06-.57i	-.06+.57i	-.53	
Inverted MA Roots	.99			

Abbildung 28: Modellierung des ARMA-Prozesses für die zweite Differenz der Spotpreise

Die zweite Differenz der Dieselpreise wird mit dem ARMA-Modell (2,3,1) approximiert. Aus der Abbildung 28 ist ersichtlich, dass das ARMA-Modell (2,3,1) für die Beschreibung der zweiten Differenzen der Dieselpreise signifikant ist, denn die p-Werte von AR(1), AR(2), AR(3), MA(1) und C nah an 0 sind. Durbin-Watson-Stat mit dem Wert von 2,016 ist nah an 2.

In der Abbildung 29 sieht man eine Übereinstimmung zwischen dem Korrelogramm für ARMA-Modell(2,3,1) und dem Korrelogramm für die zweiten Differenzen der Dieselpreise. Dies bedeutet, es liegt ein hoch signifikantes Modell vor.

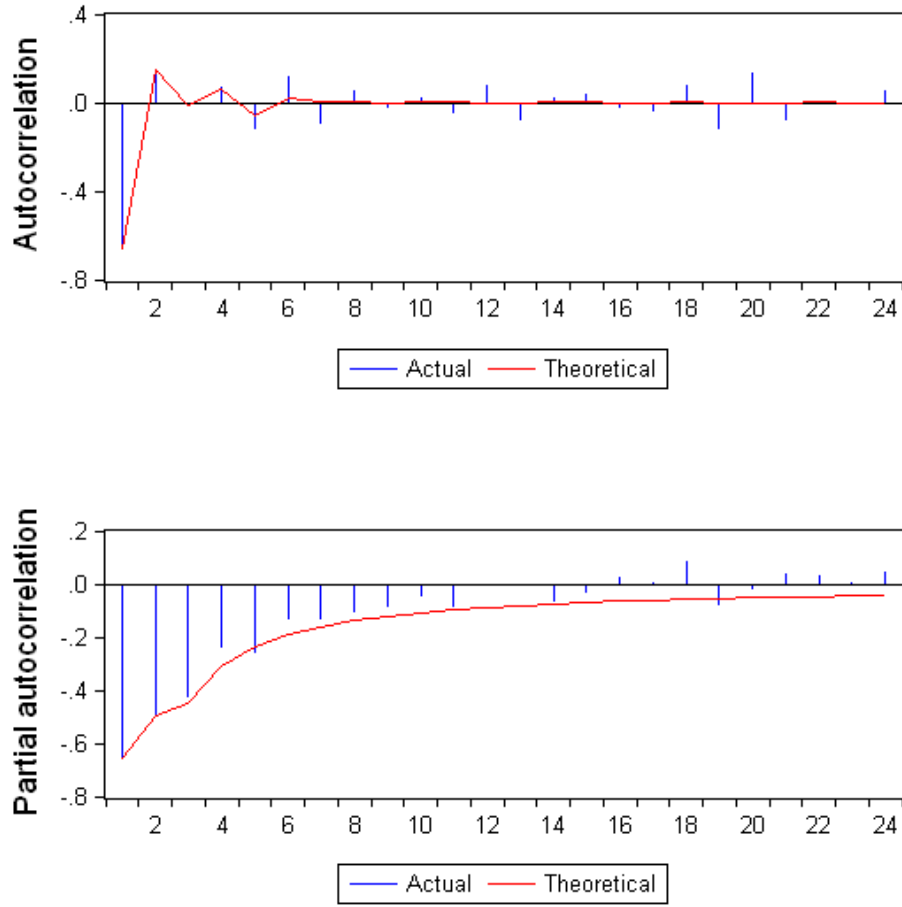


Abbildung 29: Autokorrelogramm der zweiten Differenzen der Dieselpreise

Die Zeitreihe der zweiten Differenzen der Dieselpreise wird als  $Z_t$  bezeichnet. Das gebildete ARMA-Modell(2,3,1) für diese Zeitreihe sieht folgenderweise aus:

$$Z_t = \mu + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \beta_3 Z_{t-3} + \alpha \cdot \epsilon_{t-1} + \epsilon_t \quad (5.4)$$

Mittels Integralrechnung wird das Modell für die Originalreihe der Dieselpreise  $X_t$  gebildet:

$$\begin{aligned} X_t = X_4 + \sum_{\tau=5}^{t-1} ((X_4 - X_3) + \sum_{\tau=\zeta}^{t-1} (\mu + \beta_1(X_{\zeta-1} - X_{\tau-2}) - (\beta_1 - \beta_2)(X_{\zeta-2} - X_{\zeta-3}) \\ - (\beta_2 - \beta_3)(X_{\zeta-3} - X_{\zeta-4}) - \beta_3(X_{\zeta-4} - X_{\zeta-5}) + \alpha\epsilon_{\zeta-1} + \epsilon_{\zeta})) \end{aligned} \quad (5.5)$$

Man setzt  $\alpha$ ,  $\mu$ ,  $\beta_1$ ,  $\beta_2$  und  $\beta_3$  in die Formel 5.5 ein und bekommt folgendes Modell:

$$\begin{aligned}
X_t = X_4 + \sum_{\tau=5}^{t-1} ((X_4 - X_3) + \sum_{\tau=\zeta}^{t-1} (-0,0000000498 - 0,652585(X_{\zeta-1} - X_{\tau-2}) \\
- (-0,652585 + 0,393598)(X_{\zeta-2} - X_{\zeta-3}) - (-0,393598 + 0,172214)(X_{\zeta-3} - X_{\zeta-4}) \\
- (-0,172214)(X_{\zeta-4} - X_{\zeta-5}) - 0,992617\epsilon_{\zeta-1} + \epsilon_{\zeta}))
\end{aligned} \tag{5.6}$$

Dies ist das gesuchte Modell für die Originalzeitreihe der Dieselpreise, das man für die Prognose verwenden kann. Wenn man über die Beobachtungswerte der vorherigen Perioden  $t=0, 1, 2, 3, 4$  verfügt, kann man eine Prognose für die Werte von  $X_t$  ab der Periode  $t=5$  vorhersagen.

### 5.3 Korrelation zwischen den Kraftstoffpreisen und den Spotpreisen

In der Abbildung 30 sind gleitende Durchschnitte MA6 der Kraftstoffpreise und die Originalzeitreihe der Spotpreise dargestellt. Es ist deutlich eine Gesetzmäßigkeit zu erkennen: wenn sich die Spotpreise ändern, dauert es ungefähr eine Woche, bis sich die Kraftstoffpreise ändern. Dies wird in Acht genommen und im Weiteren untersucht. Darüber hinaus merkt man, dass die Erhöhungsschritte schneller durchschlagen als die Senkungsschritte.

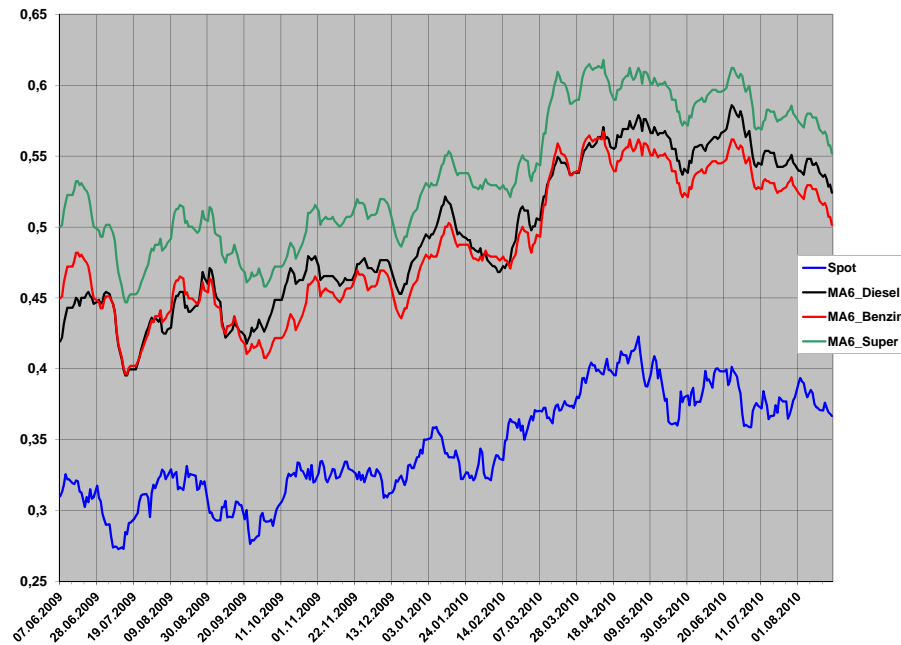


Abbildung 30: Gleitende Durchschnitte MA6 und Spotpreise

In der Abbildung 31 ist das Korrelogramm der Kreuzkorrelationsfunktion dargestellt. Es ist die Spitze an der Verzögerungsstelle 6 sehr gut sichtbar. Diese Tatsache bestätigt, dass das Verhalten der Spotpreise die Kraftstoffpreise mit einer Verzögerung von 6 Tagen beeinflusst. Da der Samstag nicht einbezogen wurde, entspricht das einer Woche.

Es wird das Regressionsmodell für die Diesel- und die Spotpreise gebildet, siehe Abbildung 32. Die Zeitreihe der Spotpreise ist für genau eine Woche vorgerückt. Die Ergebnisse des Modells sind der Abbildung 33 zu entnehmen.

In der Spalte **t-statistic** wird für die Spotpreise der Wert 44,24 angegeben. Es handelt sich um den Wert der Teststatistik des t-Tests, mit dem die Nullhypothese  $H_0: \beta = 0$  überprüft wird. Der zugehörige p-Wert wird in der Spalte *Prob.* mit 0.0000, er gibt die Wahrscheinlichkeit an, einen Wert mit einem Absolutbetrag von mindestens 44,24 zu erhalten, wenn die Nullhypothese zutrifft. Das Zutreffen der Nullhypothese ist hier also extrem unwahrscheinlich; man wird sie verwerfen und es als zutreffend ansehen, dass die Spotpreise ein Erklärungspotential für die Beobachtungen der Dieselpreise haben (Hackl, 2005).

Für die Größe **R-squared** ist der Wert 0,8388 angegeben. Es handelt sich dabei um das sogenannte Bestimmtheitsmaß, dass man als das Quadrat des Korrelationskoeffizienten zwischen den beobachteten Werten und den Schätzwerten, die die angepasste Regressionsbeziehung gibt, bzw. als Quadrat des Korrelationskoeffizienten zwischen den Spot- und Dieselpreisen interpretieren kann. Das Erklärungspotential der Spotpreise für die Dieselpreise ist entsprechend hoch einzuschätzen (Hackl, 2005).

**Durbin-Watson stat** von 1,199 deutet auf eine schwach positive Autokorrelation der Residuen. Dies kann daran liegen, dass die steigenden Spotpreise generell schneller auf die Dieselpreise durchschlagen, dabei ist der Zeitraum kleiner als der sonst von einer Woche. Die fallenden Spotpreise wirken sich aber langsamer auf die Dieselpreise aus, der Zeitraum ist größer als der sonst von einer Woche (Bundeskartellamt, 2011).

Demzufolge wurde ein Modell generiert, das zeigt, dass die Spotpreise auf die Benzinpreise mit einer Verzögerung von etwa einer Woche durchschlagen. Dies bestätigt die zweite Hypothese.

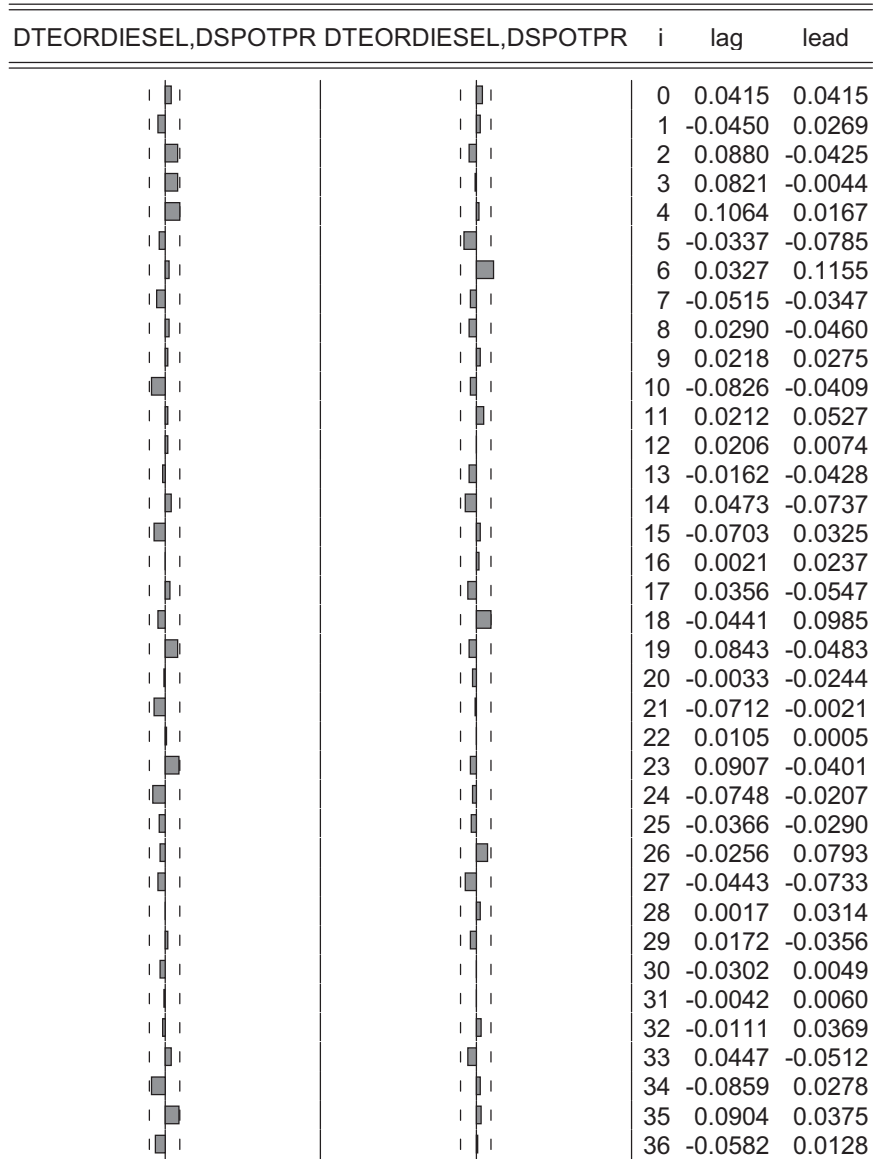


Abbildung 31: Kreuzkorrelationsfunktion

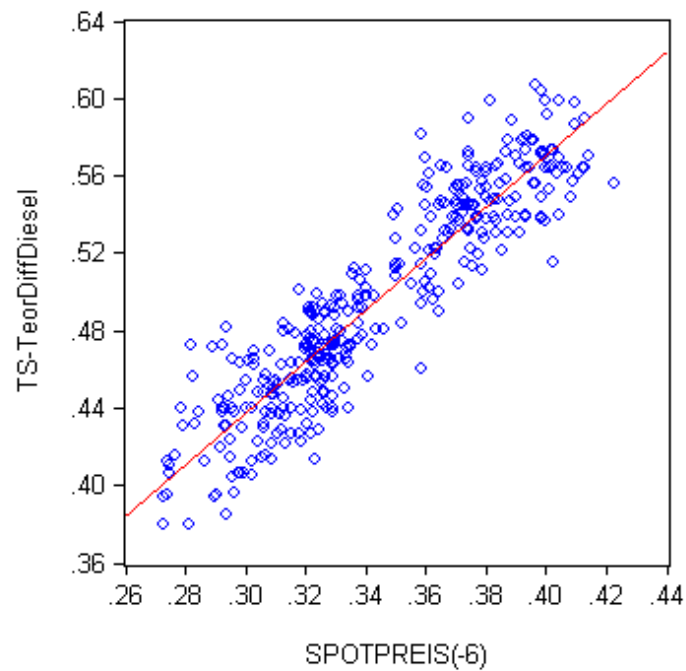


Abbildung 32: Scatterplot

Dependent Variable: TS_TEORDIFFDIESEL				
Method: Least Squares				
Date: 01/03/12 Time: 20:40				
Sample (adjusted): 6/07/2009 8/20/2010				
Included observations: 378 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.037043	0.010412	3.557547	0.0004
SPOTPREIS(-6)	1.333410	0.030138	44.24308	0.0000
R-squared	0.838865	Mean dependent var		0.495140
Adjusted R-squared	0.838437	S.D. dependent var		0.053243
S.E. of regression	0.021401	Akaike info criterion		-4.845466
Sum squared resid	0.172212	Schwarz criterion		-4.824646
Log likelihood	917.7930	Hannan-Quinn criter.		-4.837203
F-statistic	1957.450	Durbin-Watson stat		1.199158
Prob(F-statistic)	0.000000			

Abbildung 33: Regressionsmodell



## 6 Zusammenfassung

Viele empirische Untersuchungen basieren auf der Analyse von Zeitreihen. In dieser Arbeit wurde das Verhalten von Spot- und Kraftstoffpreisen mittels Zeitreihenanalyse untersucht. Die erhobenen Spot-, Diesel-, Benzin- sowie SuperPluspreise umfassen den Zeitraum von 31.05.2009 bis 20.08.2010. Die Vorbereitung und die Bereinigung der Daten forderte einen wesentlichen Zeitaufwand. Die vorläufige Analyse hat ergeben, dass alle drei Kraftstoffpreise den Verlauf voneinander wiederholen. Aus diesem Grund kann für eine zeitaufwendige Analyse eine Sorte der Kraftstoffe untersucht werden, die ein gültiges Ergebnis für alle Kraftstoffpreise liefert.

Am Anfang werden die Einheitswurzel-Tests, der erweiterte Dickey-Fuller-Test und der Kwiatkowski-Phillips-Schmidt-Shin-Test auf die Spotpreise und die Kraftstoffpreise am Beispiel der Benzinpreise angewendet. Die beiden Tests zeigen, dass die Originalzeitreihen der Benzin- und Spotpreise Nicht-Stationarität aufweisen. Deswegen werden die ersten Differenzen der Zeitreihen auf Stationarität geprüft. Die ersten Differenzen der Zeitreihen weisen nach der Durchführung des erweiterten Dickey-Fuller-Tests und des Kwiatkowski-Phillips-Schmidt-Shin-Tests Stationarität auf. Diese Eigenschaft der Zeitreihe wird einer aus statistischer Sicht guten Prognose vorausgesetzt.

Für die Spot- und Kraftstoffpreise erfolgt eine weitere Analyse auf Periodizität. Das Korrelogramm für die ersten Differenzen der Spotpreise weist keine klare Periodizität auf, da die Werte der Spitzen an den Verzögerungsstellen 48 und 75 in den Grenzen des Standardfehlers sind. Dies lässt jegliche Periodizität der Spotpreise bezweifeln. Die Korrelogramme für die ersten Differenzen aller Kraftstoffpreise zeigten dagegen große Spitzen an den durch 6 teilbaren Verzögerungsstellen. Dies deutete von einer wöchentlichen Periodizität. Genau dieses Problem hatte Enders (2010) im Zusammenhang mit dem Dickey-Fuller Test angesprochen: „Es gibt Zeitreihen, die erste Differenzierung und auch solche, die eine saisonale Differenzierung benötigen. Man muss eine Methode entwickeln, die zwischen den beiden Arten von Einheitswurzelprozessen unterscheiden kann.“

Wenn die Korrelogramme große Spitzen aufweisen, die die kleinen dominieren können, können die Ergebnisse des Modells dadurch beeinflusst sein. Deswegen wird die saisonale Zerlegung der Dieselpreise durchgeführt. Im ersten Schritt werden gleitende Durchschnitte für die Zeitreihe der Dieselpreise generiert und in beiden Modellen - additivem und multiplikativem - getestet. Es wird nicht einfach zwischen den beiden Modellen zu entscheiden. Zum Vorschein kommt das Problem des relativ kurzen Zeitraumes, der sich als nicht ausreichend für eine detaillierte saisonale Analyse erweist. Die beiden Modelle zeigen aber, dass der Freitagspreis im Wochenverlauf am höchsten ist. Dies bestätigt die erste Hypothese. Im Weiteren wird das additive Modell aufgrund deutlicherer saisonaler Abhängigkeiten bevorzugt. Die saisonale Komponente wird mittels eines Modells mit Dummy-Variablen ermittelt und eliminiert. Damit ist eine weitere Voraussetzung für die Modellbildung erfüllt. Die neue saisonbereinigte Zeitreihe wird

im Regressionsmodell angewendet.

Der Großteil der Arbeit liegt in der Modellbildung für die Diesel- und Spotpreise mit der Software Eviews. Es wird festgestellt, dass die Spotpreise durch den MA-Prozess der 1. Ordnung und die Dieselpreise durch den ARMA-Prozess (2,3,1) gut beschrieben werden. Mithilfe der Kreuzkorrelationsfunktion wird deutlich, dass das Verhalten der Dieselpreise das Verhalten der Spotpreise wiederholt. Das Regressionsmodell der Dieselpreise und der Spotpreise mit einer vorgerückten Woche bestätigt, dass die Spotpreise auf die Kraftstoffpreise in einem Zeitabstand von ca. einer Woche durchschlagen. Die zweite Hypothese wird bestätigt.

## 7 Literaturverzeichnis

- [1] Agung G (2010), *Time series data analyses using EViews*
- [2] Assenmacher W (2002), *Einführung in die Ökonometrie*, Oldenbourg Verlag
- [3] Barsky R, Miron J (1989), *The Seasonal Cycle and the Business Cycle*, Journal of Political Economy
- [4] Bundeskartellamt (2011), *Abschlussbericht zur Sektoruntersuchung der Kraftstoffe*
- [5] Dudenhöfer F (2011), *Lahme Erkenntnis vom Kartellamt*, Wirtschaftswoche, 23.05.2011
- [6] Enders W (2010), *Applied econometric time series*, New York
- [7] Energieportal, <http://www.energiesparer24.net/Sprit/Informationen/>
- [8] Fuller W (1976), *Introduction to Statistical Time Series*, New York
- [9] Hackl P (2005), *Einführung in die Ökonometrie*, Pearson Studium
- [10] Herwatz H (1995), *Analyse saisonaler Zeitreihen mit Hilfe periodischer Zeitreihenmodelle*, Verlag Josef Eul
- [11] Holding Company „Finam“, <http://www.finam.ru/analysis/export>
- [12] Kendall M (1971), *Time series*, London
- [13] Schlittgen R, Streitberg B (1989), *Zeitreihenanalyse*, Oldenbourg Verlag

- [14] Schwert G (1989), *Tests for Unit Roots: A Monte Carlo Investigation*, Journal of Business and Economic Statistics
- [15] Shumway R, Stoffer D (2000), *Time Series Analyses and Its Applications*, Springer Verlag
- [16] Sims C (1974), *Seasonality in Regression*, Journal of the American Statistical Association
- [17] Weigend A, Gershenfeld N (1994), *Time Series Prediction: Forecasting the Future and Understanding the Past*
- [18] Wettbewerbskommission (2008), *Gutachten der Wettbewerbskommission gemäß § 16 Abs. 1 Wettbewerbsgesetz an den Bundesminister für Wirtschaft und Arbeit*, Wien

## 8 Anhang

Augmented Dickey-Fuller Unit Root Test on BENZIN

Null Hypothesis: BENZIN has a unit root Exogenous: Constant, Linear Trend Lag Length: 12 (Automatic - based on SIC, maxlag=16)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.853337	0.1795
Test critical values:	1% level		-3.988333	
	5% level		-3.424579	
	10% level		-3.135349	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(BENZIN) Method: Least Squares Date: 11/22/11 Time: 17:19 Sample (adjusted): 6/23/2009 8/20/2010 Included observations: 306 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
BENZIN(-1)	-0.166658	0.058408	-2.853337	0.0046
D(BENZIN(-1))	-0.511941	0.075649	-6.767320	0.0000
D(BENZIN(-2))	-0.370748	0.078806	-4.704568	0.0000
D(BENZIN(-3))	-0.282126	0.080090	-3.522616	0.0005
D(BENZIN(-4))	-0.189083	0.080454	-2.350197	0.0194
D(BENZIN(-5))	-0.186902	0.080536	-2.320723	0.0210
D(BENZIN(-6))	0.176904	0.081901	2.159985	0.0316
D(BENZIN(-7))	0.087434	0.082566	1.058961	0.2905
D(BENZIN(-8))	0.009391	0.080994	0.115945	0.9078
D(BENZIN(-9))	-0.070570	0.078782	-0.895764	0.3711
D(BENZIN(-10))	-0.128564	0.074320	-1.729861	0.0847
D(BENZIN(-11))	-0.151965	0.068002	-2.234715	0.0262
D(BENZIN(-12))	0.183284	0.057812	3.170350	0.0017
C	0.068249	0.023999	2.843815	0.0048
@TREND(5/31/2009)	6.76E-05	2.60E-05	2.602199	0.0097
R-squared	0.605944	Mean dependent var		0.000357
Adjusted R-squared	0.586986	S.D. dependent var		0.032919
S.E. of regression	0.021156	Akaike info criterion		-4.826029
Sum squared resid	0.130242	Schwarz criterion		-4.643500
Log likelihood	753.3825	Hannan-Quinn criter.		-4.753030
F-statistic	31.96241	Durbin-Watson stat		2.002246
Prob(F-statistic)	0.000000			

Abbildung 34: Augmented Dickey-Fuller Unit Root Test für den Benzinpreis

# Augmented Dickey-Fuller Unit Root Test on SPOT

Null Hypothesis: SPOT has a unit root Exogenous: Constant, Linear Trend Lag Length: 0 (Automatic - based on SIC, maxlag=16)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-2.754462	0.2155
Test critical values:	1% level		-3.983684	
	5% level		-3.422321	
	10% level		-3.134016	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(SPOT) Method: Least Squares Date: 11/22/11 Time: 18:28 Sample (adjusted): 6/01/2009 8/20/2010 Included observations: 361 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
SPOT(-1)	-0.044055	0.015994	-2.754462	0.0062
C	0.013135	0.004670	2.812670	0.0052
@TREND(5/31/2009)	1.09E-05	5.28E-06	2.068606	0.0393
R-squared	0.021645	Mean dependent var		7.96E-05
Adjusted R-squared	0.016179	S.D. dependent var		0.005780
S.E. of regression	0.005733	Akaike info criterion		-7.476956
Sum squared resid	0.011766	Schwarz criterion		-7.444638
Log likelihood	1352.591	Hannan-Quinn criter.		-7.464107
F-statistic	3.960123	Durbin-Watson stat		2.019651
Prob(F-statistic)	0.019902			

Abbildung 35: Augmented Dickey-Fuller Unit Root Test für den Spotpreis

Augmented Dickey-Fuller Unit Root Test on D(BENZIN)

Null Hypothesis: D(BENZIN) has a unit root Exogenous: Constant, Linear Trend Lag Length: 11 (Automatic - based on SIC, maxlag=16)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-6.275240	0.0000
Test critical values:	1% level		-3.988333	
	5% level		-3.424579	
	10% level		-3.135349	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(BENZIN,2) Method: Least Squares Date: 11/22/11 Time: 20:07 Sample (adjusted): 6/23/2009 8/20/2010 Included observations: 306 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(BENZIN(-1))	-3.324279	0.529745	-6.275240	0.0000
D(BENZIN(-1),2)	1.670189	0.506793	3.295600	0.0011
D(BENZIN(-2),2)	1.179644	0.478037	2.467684	0.0142
D(BENZIN(-3),2)	0.794662	0.446459	1.779921	0.0761
D(BENZIN(-4),2)	0.517020	0.412700	1.252776	0.2113
D(BENZIN(-5),2)	0.252132	0.375951	0.670652	0.5030
D(BENZIN(-6),2)	0.361390	0.332727	1.086143	0.2783
D(BENZIN(-7),2)	0.377572	0.278797	1.354287	0.1767
D(BENZIN(-8),2)	0.320216	0.224326	1.427461	0.1545
D(BENZIN(-9),2)	0.192189	0.167812	1.145261	0.2530
D(BENZIN(-10),2)	0.017734	0.112041	0.158280	0.8743
D(BENZIN(-11),2)	-0.167411	0.058243	-2.874348	0.0043
C	0.000156	0.002570	0.060602	0.9517
@TREND(5/31/2009)	1.07E-06	1.16E-05	0.091851	0.9269
R-squared	0.862192	Mean dependent var		0.000110
Adjusted R-squared	0.856056	S.D. dependent var		0.056439
S.E. of regression	0.021413	Akaike info criterion		-4.804972
Sum squared resid	0.133886	Schwarz criterion		-4.634611
Log likelihood	749.1606	Hannan-Quinn criter.		-4.736839
F-statistic	140.5297	Durbin-Watson stat		1.993761
Prob(F-statistic)	0.000000			

Abbildung 36: Augmented Dickey-Fuller Unit Root Test für den Benzinpreis. Erste Differenz

# Augmented Dickey-Fuller Unit Root Test on D(SPOT)

Null Hypothesis: D(SPOT) has a unit root Exogenous: Constant, Linear Trend Lag Length: 0 (Automatic - based on SIC, maxlag=16)				
			t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic			-18.96369	0.0000
Test critical values:	1% level		-3.984496	
	5% level		-3.422716	
	10% level		-3.134249	
*MacKinnon (1996) one-sided p-values.				
Augmented Dickey-Fuller Test Equation Dependent Variable: D(SPOT,2) Method: Least Squares Date: 11/22/11 Time: 20:11 Sample (adjusted): 6/02/2009 8/20/2010 Included observations: 350 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SPOT(-1))	-1.022017	0.053893	-18.96369	0.0000
C	0.000439	0.000620	0.709095	0.4787
@TREND(5/31/2009)	-1.99E-06	2.81E-06	-0.710088	0.4781
R-squared	0.508932	Mean dependent var		-3.17E-05
Adjusted R-squared	0.506102	S.D. dependent var		0.008317
S.E. of regression	0.005845	Akaike info criterion		-7.437977
Sum squared resid	0.011854	Schwarz criterion		-7.404909
Log likelihood	1304.646	Hannan-Quinn criter.		-7.424815
F-statistic	179.8116	Durbin-Watson stat		2.016351
Prob(F-statistic)	0.000000			

Abbildung 37: Augmented Dickey-Fuller Unit Root Test für den Spotpreis. Erste Differenz



KPSS Unit Root Test on BENZIN

Null Hypothesis: BENZIN is stationary Exogenous: Constant, Linear Trend Bandwidth: 14 (Newey-West automatic) using Bartlett kernel				
				LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic				0.219196
Asymptotic critical values*:				0.216000
1% level				0.146000
5% level				0.119000
10% level				
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)				
Residual variance (no correction)				0.001105
HAC corrected variance (Bartlett kernel)				0.007723
KPSS Test Equation Dependent Variable: BENZIN Method: Least Squares Date: 11/24/11 Time: 17:32 Sample: 5/31/2009 8/20/2010 Included observations: 376				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.416580	0.003452	120.6831	0.0000
@TREND(5/31/2009)	0.000365	1.56E-05	23.44339	0.0000
R-squared	0.595059	Mean dependent var		0.486757
Adjusted R-squared	0.593977	S.D. dependent var		0.052306
S.E. of regression	0.033330	Akaike info criterion		-3.959439
Sum squared resid	0.415462	Schwarz criterion		-3.938537
Log likelihood	746.3745	Hannan-Quinn criter.		-3.951142
F-statistic	549.5923	Durbin-Watson stat		0.924051
Prob(F-statistic)	0.000000			

Abbildung 38: KPSS-Test für den Benzinpreis.

KPSS Unit Root Test on D(BENZIN)

Null Hypothesis: D(BENZIN) is stationary				
Exogenous: Constant, Linear Trend				
Bandwidth: 31 (Newey-West automatic) using Bartlett kernel				
				LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic				0.050980
Asymptotic critical values*:				
1% level				0.216000
5% level				0.146000
10% level				0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)				
Residual variance (no correction)				0.001024
HAC corrected variance (Bartlett kernel)				8.42E-05
KPSS Test Equation Dependent Variable: D(BENZIN) Method: Least Squares Date: 11/24/11 Time: 16:33 Sample (adjusted): 6/01/2009 8/20/2010 Included observations: 370 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001011	0.003368	0.300186	0.7642
@TREND(5/31/2009)	-3.24E-06	1.52E-05	-0.213573	0.8310
R-squared	0.000124	Mean dependent var		0.000386
Adjusted R-squared	-0.002593	S.D. dependent var		0.032040
S.E. of regression	0.032082	Akaike info criterion		-4.035680
Sum squared resid	0.378755	Schwarz criterion		-4.014526
Log likelihood	748.6008	Hannan-Quinn criter.		-4.027277
F-statistic	0.045613	Durbin-Watson stat		2.929652
Prob(F-statistic)	0.830998			

Abbildung 39: KPSS-Test für den Benzinpreis. Erste Differenz

KPSS Unit Root Test on SPOT

Null Hypothesis: SPOT is stationary Exogenous: Constant, Linear Trend Bandwidth: 15 (Newey-West automatic) using Bartlett kernel				
				LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic				0.166831
Asymptotic critical values*:				
1% level				0.216000
5% level				0.146000
10% level				0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)				
Residual variance (no correction)				0.000352
HAC corrected variance (Bartlett kernel)				0.004176
KPSS Test Equation Dependent Variable: SPOT Method: Least Squares Date: 11/30/11 Time: 14:11 Sample: 5/31/2009 8/20/2010 Included observations: 372				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.290016	0.001936	149.7707	0.0000
@TREND(5/31/2009)	0.000282	8.76E-06	32.21098	0.0000
R-squared	0.737131	Mean dependent var		0.343889
Adjusted R-squared	0.736421	S.D. dependent var		0.036662
S.E. of regression	0.018822	Akaike info criterion		-5.102198
Sum squared resid	0.131082	Schwarz criterion		-5.081129
Log likelihood	951.0089	Hannan-Quinn criter.		-5.093831
F-statistic	1037.547	Durbin-Watson stat		0.094394
Prob(F-statistic)	0.000000			

Abbildung 40: KPSS-Test für den Spotpreis.

KPSS Unit Root Test on D(SPOT)

Null Hypothesis: D(SPOT) is stationary Exogenous: Constant, Linear Trend Bandwidth: 8 (Newey-West automatic) using Bartlett kernel				
				LM-Stat.
Kwiatkowski-Phillips-Schmidt-Shin test statistic				0.048129
Asymptotic critical values*:				1% level 0.216000
				5% level 0.146000
				10% level 0.119000
*Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1)				
Residual variance (no correction)				3.33E-05
HAC corrected variance (Bartlett kernel)				2.86E-05
KPSS Test Equation Dependent Variable: D(SPOT) Method: Least Squares Date: 11/30/11 Time: 14:13 Sample (adjusted): 6/01/2009 8/20/2010 Included observations: 361 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000378	0.000604	0.625741	0.5319
@TREND(5/31/2009)	-1.56E-06	2.73E-06	-0.571999	0.5677
R-squared	0.000911	Mean dependent var		7.96E-05
Adjusted R-squared	-0.001872	S.D. dependent var		0.005780
S.E. of regression	0.005785	Akaike info criterion		-7.461525
Sum squared resid	0.012015	Schwarz criterion		-7.439980
Log likelihood	1348.805	Hannan-Quinn criter.		-7.452959
F-statistic	0.327183	Durbin-Watson stat		2.066421
Prob(F-statistic)	0.567681			

Abbildung 41: KPSS-Test für den Spotpreis. Erste Differenz

## Erklärung zur Urheberschaft

Hiermit versichere ich, dass ich die vorliegende Arbeit zum Thema *Anwendung der Methoden und Modelle der Zeitreihenanalyse auf Spot- und Kraftstoffpreise* selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel verfasst habe. Die Prüfungsordnung ist mir bekannt. Ich habe in meinem Studienfach bisher keine Bachelorarbeit eingereicht bzw. sie nicht endgültig nicht bestanden.

Iryna Karpenka

8. Januar 2012